The Evidence Interface

How policymakers encounter, engage with, and make sense of scientific knowledge

Sebastian Ramirez-Ruiz

(Data Science Lab — Hertie School)

Dissertation Defense

November 11, 2025

Hertie School

Evidence can be an important **input** in **policymaking**

Politicians and policymakers often **invoke** the **authority** of scientific knowledge and technical expertise

Evidence can be an important **input** in **policymaking**

Politicians and policymakers often **invoke** the **authority** of scientific knowledge and technical expertise

- food standards
- public health
- climate response
- …

Evidence can be an important **input** in **policymaking**

Politicians and policymakers often **invoke** the **authority** of scientific knowledge and technical expertise

- food standards
- public health
- climate response
- …

Evidence can be an important **input** in **policymaking**

Narrative spread of **evidence-based policymaking** (*and praxis?*) across the world

Politicians and policymakers often **invoke** the **authority** of scientific knowledge and technical expertise

- food standards
- public health
- climate response
- …

Evidence can be an important **input** in **policymaking**

Formalization of governance arrangements and *grand strategies* for **evidence integration** across polities (e.g., U.S., U.K., EU, AU)

Narrative spread of **evidence-based policymaking** (*and praxis?*) across the world

Despite its rational appeal, the notion of policymaking **guided** by **evidence** leaves some open questions

Despite its rational appeal, the notion of policymaking **guided** by **evidence** leaves some open questions

What counts as **evidence**?

What counts as **evidence**?

Whose expertise permeates decisions?

What counts as **evidence**?

Whose expertise permeates decisions?

How does knowledge flow—*if at all*—from research and expert communities to elite decision-makers?

What counts as **evidence**?

Whose expertise permeates decisions?

How does knowledge flow—*if at all*—from research and expert communities to elite decision-makers?

And **do policymakers have** the **competencies** to work with evidence in the first place?

(e.g., Caplan, 1979; Weiss, 1979; Head, 2008; Contandriopoulos et al., 2010; National Research Council, 2012; Parkhurst, 2017; Walgrave and Dejaeghere, 2017; Senninger and Hansen, 2025)

In practice, information and evidence flows are:

(e.g., Caplan, 1979; Weiss, 1979; Head, 2008; Contandriopoulos et al., 2010; National Research Council, 2012; Parkhurst, 2017; Walgrave and Dejaeghere, 2017; Senninger and Hansen, 2025)

In practice, information and evidence flows are:

fragmented,

(e.g., Caplan, 1979; Weiss, 1979; Head, 2008; Contandriopoulos et al., 2010; National Research Council, 2012; Parkhurst, 2017; Walgrave and Dejaeghere, 2017; Senninger and Hansen, 2025)

In practice, information and evidence flows are:

fragmented,

contested,

(e.g., Caplan, 1979; Weiss, 1979; Head, 2008; Contandriopoulos et al., 2010; National Research Council, 2012; Parkhurst, 2017; Walgrave and Dejaeghere, 2017; Senninger and Hansen, 2025)

In practice, information and evidence flows are:

fragmented,
contested,
selective

(e.g., Caplan, 1979; Weiss, 1979; Head, 2008; Contandriopoulos et al., 2010; National Research Council, 2012; Parkhurst, 2017; Walgrave and Dejaeghere, 2017; Senninger and Hansen, 2025)

Studying **elites** is **not** an **easy** feat

Studying **‘the evidence interface’** presents some challenges

Studying **'the evidence interface'** presents some challenges:

- Interactions happen behind closed doors or through informal routines.

Studying **'the evidence interface'** presents some challenges:

- Interactions happen behind closed doors or through informal routines.
- Common methods (interviews, surveys, information provision experiments) provide insight but:

Studying **'the evidence interface'** presents some challenges:

- Interactions happen behind closed doors or through informal routines.
- Common methods (interviews, surveys, information provision experiments) provide insight but:
  - rely on **retrospective** or **self-reported** accounts,

Studying **'the evidence interface'** presents some challenges:

- Interactions happen behind closed doors or through informal routines.
- Common methods (interviews, surveys, information provision experiments) provide insight but:
  - rely on **retrospective** or **self-reported** accounts,
  - are **hard to scale** across countries/time,

Studying **'the evidence interface'** presents some challenges:

- Interactions happen behind closed doors or through informal routines.
- Common methods (interviews, surveys, information provision experiments) provide insight but:
  - rely on **retrospective** or **self-reported** accounts,
  - are **hard to scale** across countries/time,
  - miss the **granular** and **fluid** dynamics of information **flows**.

Large emphasis on:

(e.g., Baekgaard et al., 2019; Banuri et al., 2019; DellaVigna et al., 2024; Hjort et al., 2021; Lee, 2022; Li, 2017; Toma and Bell, 2024; Vivalt and Coville, 2023)

Large emphasis on:

"What **would** a policy decisionmaker **do if** she were **confronted** with **X** research **finding**?"

(e.g., Baekgaard et al., 2019; Banuri et al., 2019; DellaVigna et al., 2024; Hjort et al., 2021; Lee, 2022; Li, 2017; Toma and Bell, 2024; Vivalt and Coville, 2023)

Information does not simply **“travel”** from researchers to elites—it is **filtered, selected,** and **‘digested’** within complex environments.

Individual inclination to **engage with** researchers and **self-selection** into research evidence can matter

Individual inclination to **engage with** researchers and **self-selection** into research evidence can matter

a lot

With that in mind, in the dissertation I take a **step back** and ask:

With that in mind, in the dissertation I take a **step back** and ask:

"**Would** she **encounter** researchers and be **confronted** with **X** research **finding** in the first place?"

What counts as **evidence**?

What counts as **evidence**?

Whose expertise permeates decisions?

What counts as **evidence**?

Whose expertise permeates decisions?

How does knowledge flow—*if at all*—from research and expert communities to elite decision-makers?

What counts as **evidence**?

Whose expertise permeates decisions?

How does knowledge flow—*if at all*—from research and expert communities to elite decision-makers?

And **do policymakers have** the **competencies** to work with evidence in the first place?

My dissertation contains **four individual papers**,

My dissertation contains **four individual papers**,

- Politicians from 12 countries rarely engage with researchers on social media, but this can change when expertise gains salience (Ramirez-Ruiz)
- Policy documents across 185 countries predominantly rely on evidence from the Global North (Ramirez-Ruiz and Senninger)
- The Bundestag Expert Witness Tracker (BEWiT): A database of German Bundestag public expert hearings (Ramirez-Ruiz)
- Measuring scientific evidence consumption literacy for public policy: Development and validation of the INSPIRE inventory (Ramirez-Ruiz and Munzert)

My dissertation contains **four individual papers**,

- Politicians from 12 countries rarely engage with researchers on social media, but this can change when expertise gains salience (Ramirez-Ruiz)
- Policy documents across 185 countries predominantly rely on evidence from the Global North (Ramirez-Ruiz and Senninger)
- The Bundestag Expert Witness Tracker (BEWiT): A database of German Bundestag public expert hearings (Ramirez-Ruiz)
- Measuring scientific evidence consumption literacy for public policy: Development and validation of the INSPIRE inventory (Ramirez-Ruiz and Munzert)

What counts as **evidence**?

Whose expertise permeates decisions?

How does knowledge flow—*if at all*—from research and expert communities to elite decision-makers?

- I want to know if what **we** (*scientists*) **discover** and **talk about** actually makes it in front of people who make big decisions.

- I want to know if what **we** (*scientists*) **discover** and **talk about** actually makes it in front of people who make big decisions.
- To study this, I look for clues in places people do not usually check (*like traces of stuff people and governments do or store online*)

- I want to know if what **we** (*scientists*) **discover** and **talk about** actually makes it in front of people who make big decisions.
- To study this, I look for clues in places people do not usually check (*like traces of stuff people and governments do or store online*)
- I gather lots of these clues, curate them, organize them, and try to understand the patterns.

My dissertation contains **four individual papers**,

- Politicians from 12 countries rarely engage with researchers on social media, but this can change when expertise gains salience (Ramirez-Ruiz)
- Policy documents across 185 countries predominantly rely on evidence from the Global North (Ramirez-Ruiz and Senninger)
- The Bundestag Expert Witness Tracker (BEWiT): A database of German Bundestag public expert hearings (Ramirez-Ruiz)
- Measuring scientific evidence consumption literacy for public policy: Development and validation of the INSPIRE inventory (Ramirez-Ruiz and Munzert)

What counts as **evidence**?

Whose expertise permeates decisions?

How does knowledge flow—*if at all*—from research and expert communities to elite decision-makers?

And **do policymakers have** the **competencies** to work with evidence in the first place?

individuals

Politicians from 12 countries rarely engage with researchers on social media, but this can change when expertise gains salience

Sebastian Ramirez-Ruiz (Hertie School)*

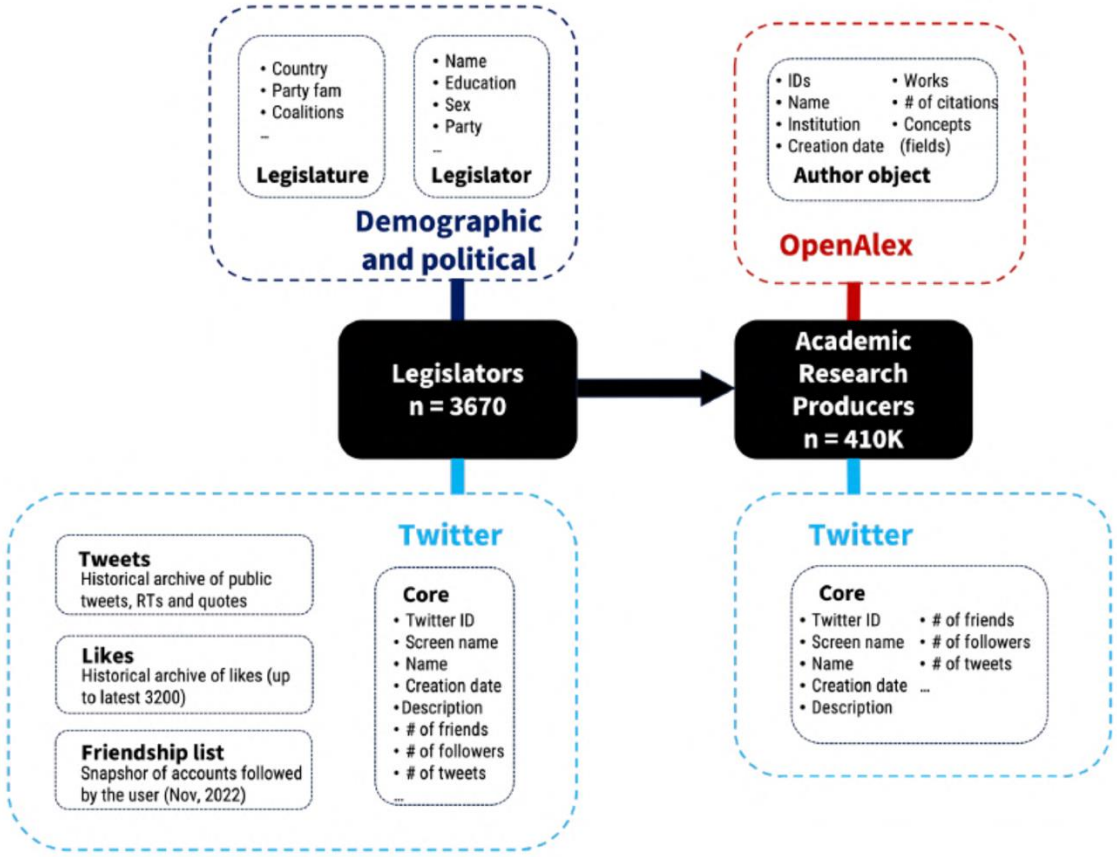
Abstract. Interactions between the policy and academic communities can play an important role in political decisionmaking. Still, the fact that much of the policymaking process happens behind closed doors obscures our understanding of the relationships between political decision-makers with academic researchers. To address this challenge, this paper introduces a novel approach that leverages online behavioral data from social media to examine how legislators interact with researchers. By analyzing data from 3,670 lawmakers in 12 countries merged to a novel database of 410K academic researchers on Twitter, this study provides new insights into these otherwise hidden interactions. The findings suggest that lawmakers do follow, yet rarely visibly engage with researchers online. Lawmakers from conservative and radical right parties follow and engage less with researchers online than their colleagues from other parties. While the base engagement is relatively low across legislatures, it can increase when expertise gains salience. During the early stages of the COVID-19 pandemic, marked by policy uncertainty involving a novel and technically complex policy issue, lawmakers' overall inclination to follow and engage with scholars increased, most prominently targeting researchers from the medical sciences. These findings offer new insights into when and how lawmakers publicly attend to academic expertise, contributing to a broader understanding of political elites' symbolic and informational engagement with science.

Keywords. legislative elites | academic researchers | elite digital traces | social media

(under review at the *British Journal of Political Science*)

In this project, I collect Twitter **digital trace data** from ≈**3.7K legislators in 12 countries** and map it onto a novel dataset of more than **410K researchers**

COUNTRY	SEATS	ON TWITTER
Pooled	**4134**	**3670**
Ireland	160	154
UK	650	590
Spain	349	314
Colombia	108	104
Canada	338	325
Argentina	72	61
Germany	736	597
US Senate	100	100
US House	435	431
Italy	400	286
Mexico	128	114
France	577	552
Brazil	81	76



In this project, I collect Twitter **digital trace data** from ≈**3.7K legislators in 12 countries** and map it onto a novel dataset of more than **410K researchers** to answer three sequential questions:

In this project, I collect Twitter **digital trace data** from **≈3.7K legislators in 12 countries** and map it onto a novel dataset of more than **410K researchers** to answer three sequential questions:

- Do we **observe** lawmakers following and engaging with researchers **'in the wild'**?

In this project, I collect Twitter **digital trace data** from **≈3.7K legislators in 12 countries** and map it onto a novel dataset of more than **410K researchers** to answer three sequential questions:

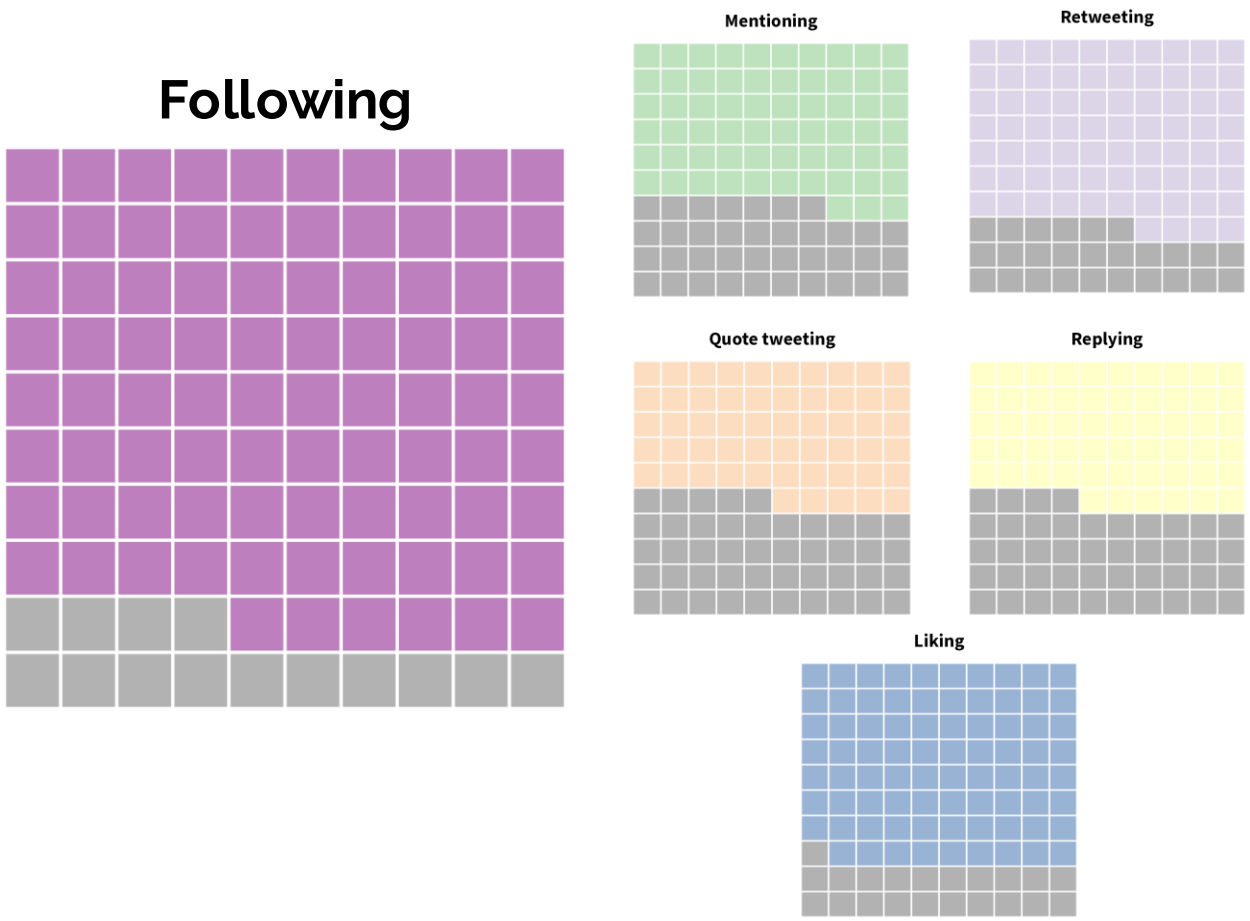
- Do we **observe** lawmakers following and engaging with researchers **'in the wild'**?
- Are there any contextual and legislator level **correlates** to online **engagement** with researchers?

In this project, I collect Twitter **digital trace data** from **≈3.7K legislators in 12 countries** and map it onto a novel dataset of more than **410K researchers** to answer three sequential questions:

- Do we **observe** lawmakers following and engaging with researchers **'in the wild'**?
- Are there any contextual and legislator level **correlates** to online **engagement** with researchers?
- Do legislators **adapt** their behaviors to **exogenous shocks** to the salience of expertise?

What did I find?

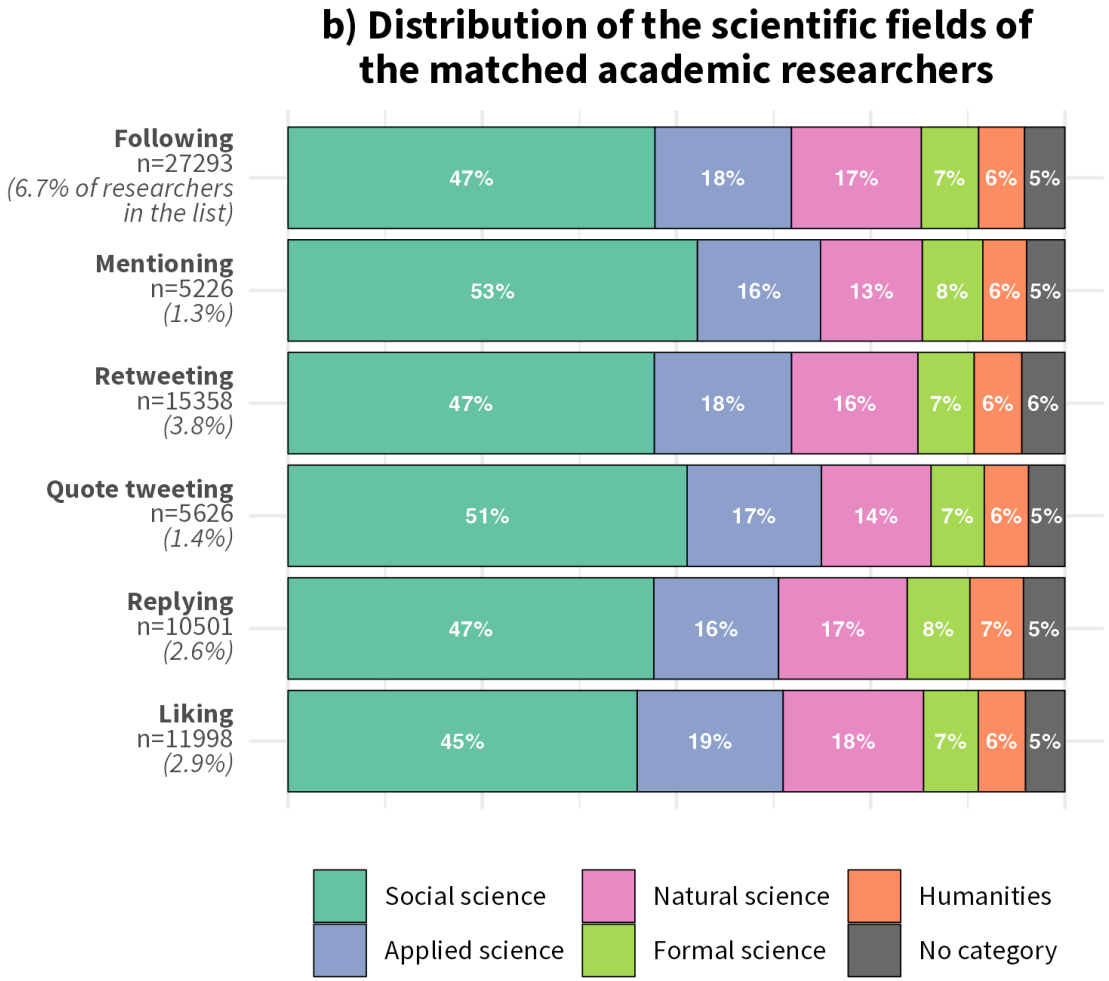
1. Do we **observe** lawmakers following and engaging with researchers **'in the wild'**?



The **majority** of these legislators do **follow** and **engage** with researchers

What did I find?

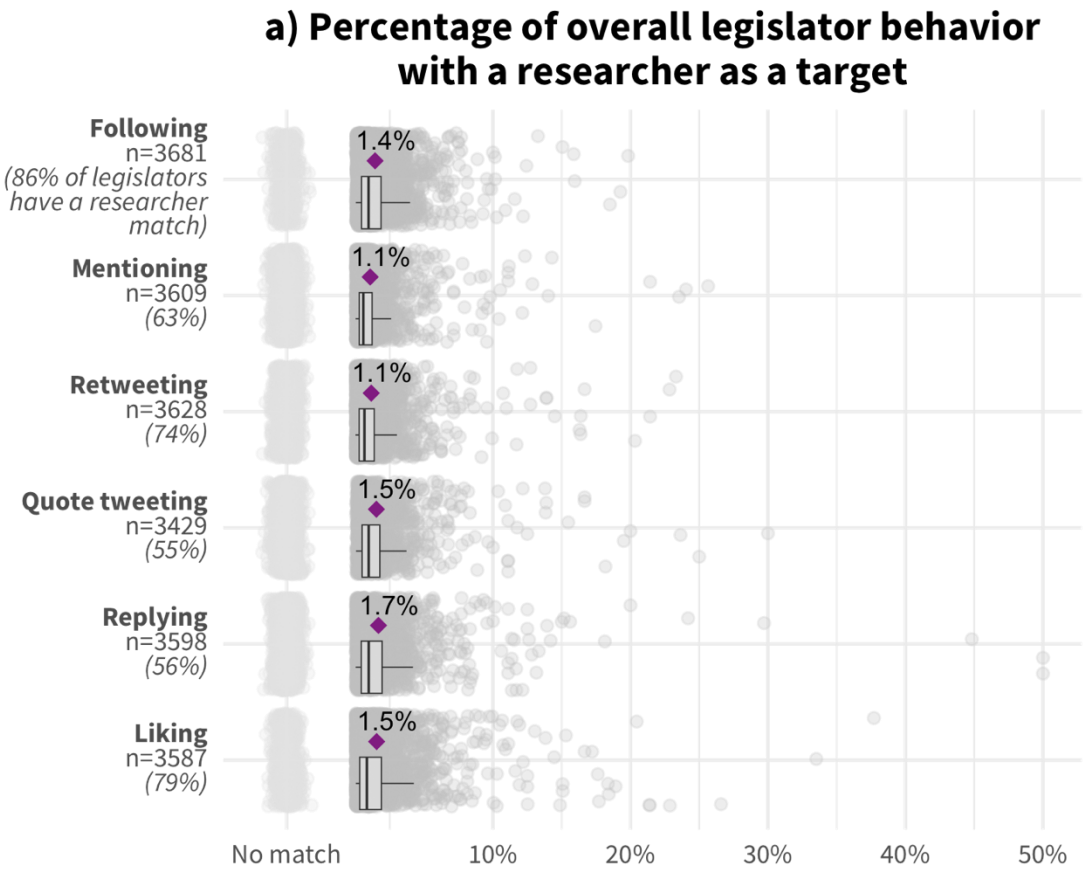
1. Do we **observe** lawmakers following and engaging with researchers **'in the wild'**?



Most legislators' "attention" goes to **social scientists**

What did I find?

1. Do we **observe** lawmakers following and engaging with researchers **'in the wild'**?

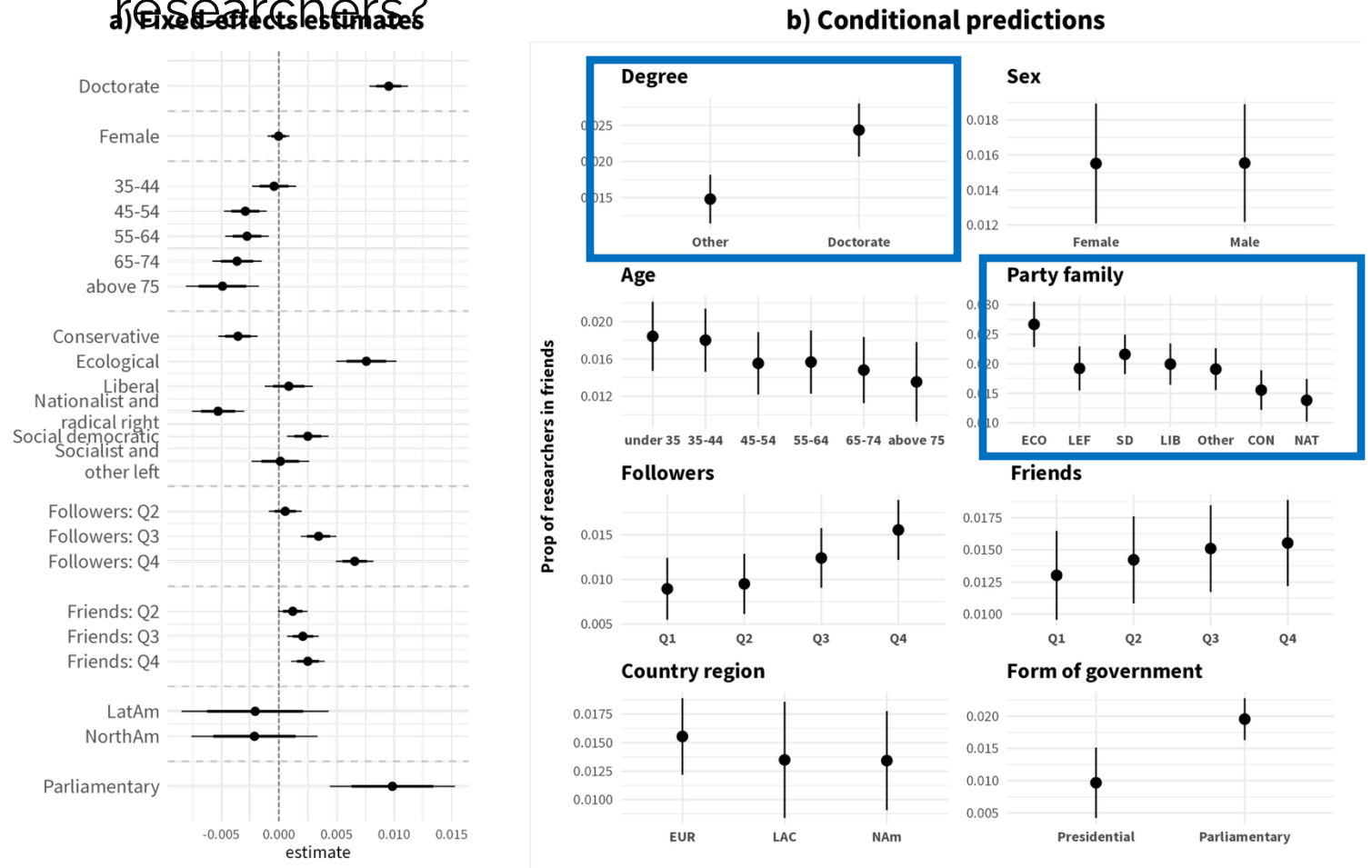


There are **some differences**, but for the most part these researchers represent a small fraction of whom lawmakers follow and engage with on social media

What did I find?

Legislators' **research background** and **political ideology** are strong predictors across behaviors

2. Are there any contextual and legislator level **correlates** to online **engagement** with researchers?



Estimated effects of legislator and legislature characteristics on the proportion of researchers in their networks. Estimated effects of legislator and legislature characteristics on the proportion of researchers in their networks.} Results from a linear mixed-effects model with legislature random effects with age (under 35), party family (other), country region (Europe), system (presidential), and Q1 for followers and friends as references for categorical variables. Number of observations: 3,247. Panel a presents the coefficients with 80% and 95% confidence intervals. The conditional predictions are computed with numeric covariates are held at their means and the other covariates at their modes: no research degree, presidential, European, male, 45-54, Q1, and Conservative party..

What did I find?

2. Are there any contextual and legislator level **correlates** to online **engagement** with researchers?

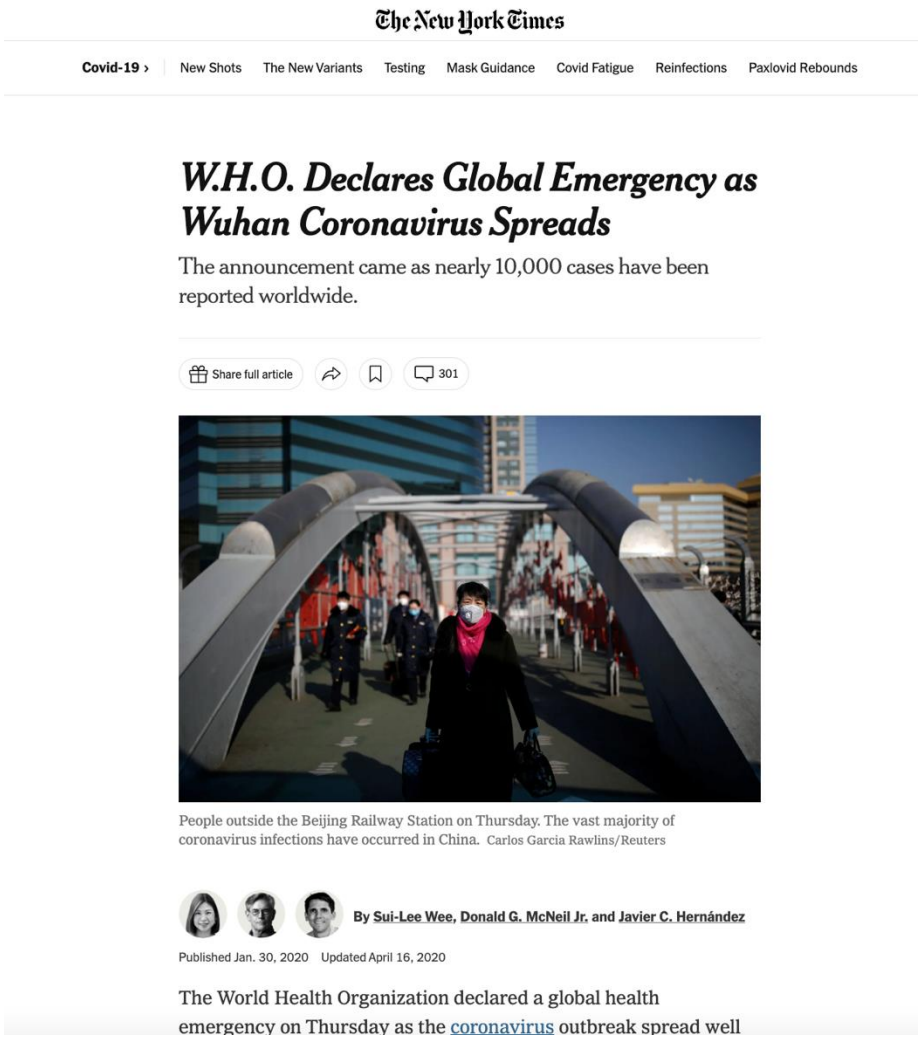
A **'green'** legislator is **3.5x** more likely to follow researchers compared to a **'nationalist and radical right'** member

Estimated effects of legislator and legislature characteristics on the proportion of researchers in their networks. Estimated effects of legislator and legislature characteristics on the proportion of researchers in their networks.} Results from a linear mixed-effects model with legislature random effects with age (under 35), party family (other), country region (Europe), system (presidential), and Q1 for followers and friends as references for categorical variables. Number of observations: 3,247. Panel a presents the coefficients with 80% and 95% confidence intervals. The conditional predictions are computed with numeric covariates are held at their means and the other covariates at their modes: no research degree, presidential, European, male, 45-54, Q1, and Conservative party..

What did I find?

3. Do legislators **adapt** their behaviors to exogenous **shocks** to the salience of **expertise**?

Legislators' digital behaviors seem to be responsive to exogenous **shocks** to the **salience of expertise**



The New York Times

Covid-19 › | New Shots | The New Variants | Testing | Mask Guidance | Covid Fatigue | Reinfections | Paxlovid Rebounds

W.H.O. Declares Global Emergency as Wuhan Coronavirus Spreads

The announcement came as nearly 10,000 cases have been reported worldwide.

Share full article | 301

People outside the Beijing Railway Station on Thursday. The vast majority of coronavirus infections have occurred in China. Carlos Garcia Rawlins/Reuters

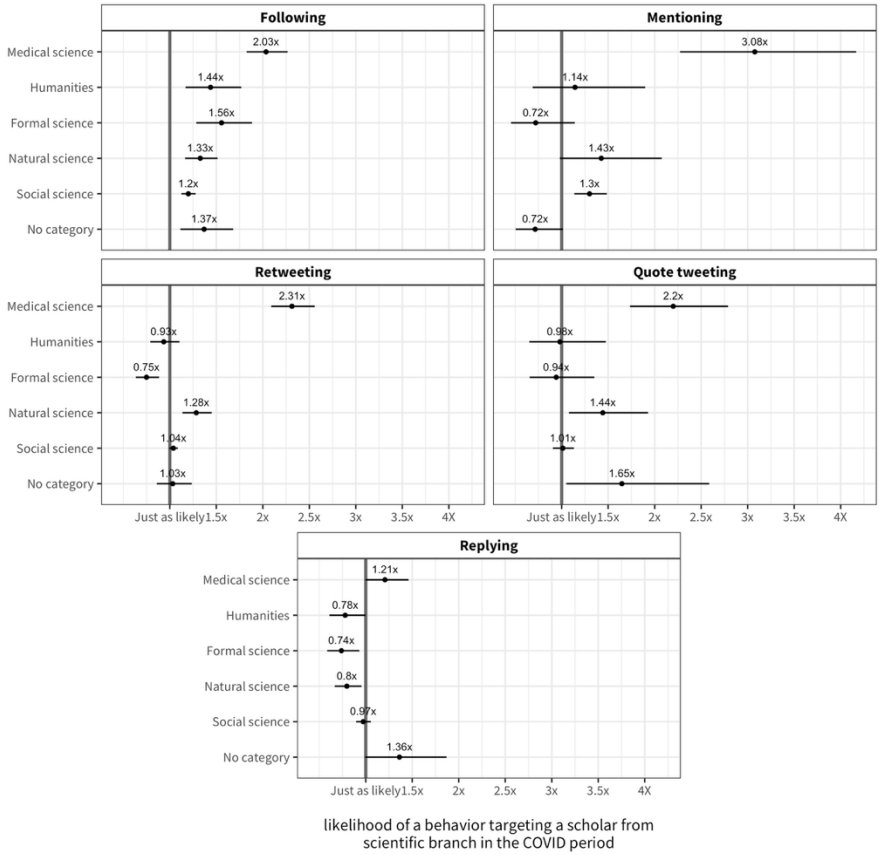
By Sui-Lee Wee, Donald G. McNeil Jr. and Javier C. Hernández

Published Jan. 30, 2020 Updated April 16, 2020

The World Health Organization declared a global health emergency on Thursday as the coronavirus outbreak spread well

What did I find?

3. Do legislators **adapt** their behaviors to exogenous **shocks** to the salience of **expertise**?



Observable increases in **engagement** targeted at researchers in the **medical sciences** in the immediate aftermath of COVID-19

Marginal effects of public health crisis on following and engagement with academic researchers during the COVID versus pre-COVID periods with a ±12 week bandwidth. Results from a logistic mixed-effects models with legislature random effects. The estimates in the figure are relative risks representing the ratio of the probability of an event in the COVID period to the probability of an outcome in a pre-COVID period..

artifacts of policymaking

Policy documents across 185 countries predominantly rely on evidence from the Global North

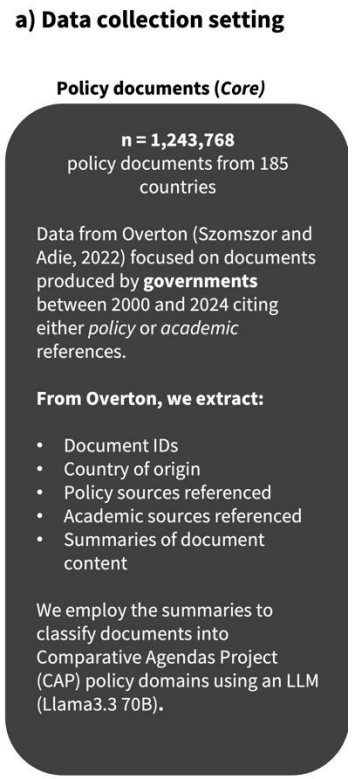
Sebastian Ramirez-Ruiz (Hertie School)*
Roman Senninger (Aarhus University)

Abstract. Evidence is widely acknowledged as essential for crafting effective public policies. Despite its critical role, we know surprisingly little about the specific sources that inform decisions around the world. This paper explores the sources of evidence in the policymaking arena by analyzing evidence cited in over 1.2 million policy documents from 185 countries. Our analyses capture references to 3.5 million scholarly works and 740,000 policy sources including contributions from government agencies, academic researchers, international organizations, and think tanks. By focusing on the *documented*, *accessible*, and *digitally visible* evidence available to policymakers, we map global patterns of evidence use, highlighting regional and policy domain variation. Our findings reveal a pronounced concentration of attention: the vast majority of cited evidence—both academic and policy—is produced in the Global North, even in documents authored by governments in the Global South. These patterns persist across policy areas, though with notable variation in the types of sources commonly used. Overall, the findings reveal a highly concentrated evidence landscape, where a small number of countries disproportionately serve as global reference points, underscoring persistent asymmetries in visibility, influence, and access within the international policy knowledge ecosystem.

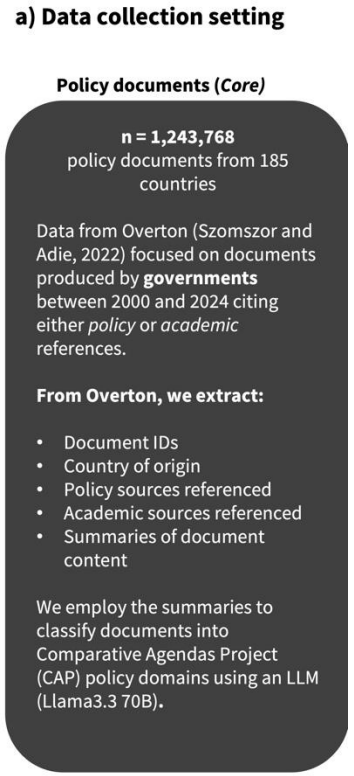
Keywords. Evidence-informed policymaking; Expert-based information; Information diffusion; Science-policy nexus; Global policy analysis; Knowledge brokers

(R&R at *Nature Human Behaviour*)

In this project, we analyze data from **1.2 million government policy documents** from **185 countries**

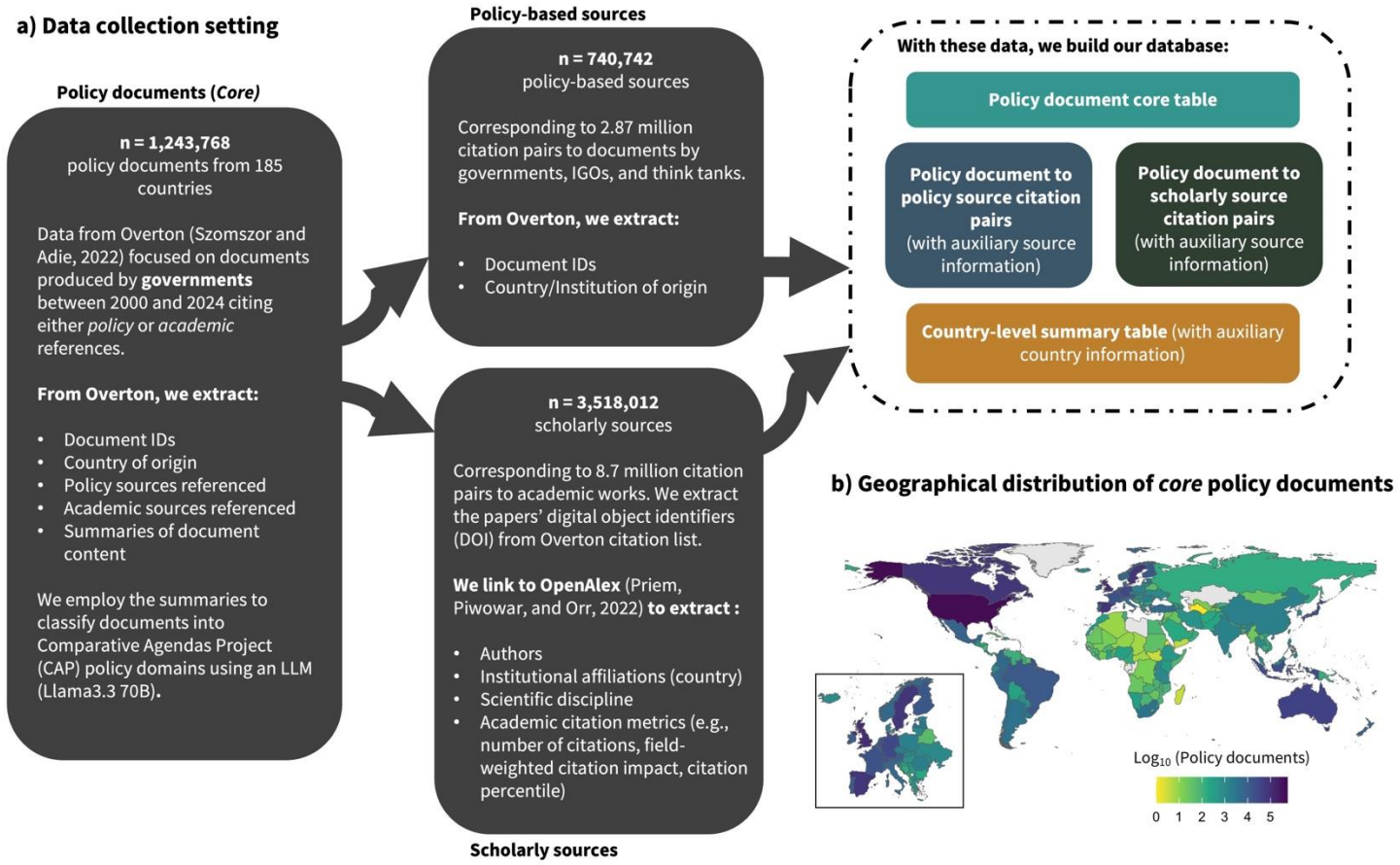


In this project, we analyze data from **1.2 million government policy documents** from **185 countries** extracting the sources they cite



- ***Policy-based*** (govs, IGOs, think tanks)
- ***Scholarly*** (journals, working papers)

In this project, we analyze data from **1.2 million government policy documents** from **185 countries** extracting the sources they cite to build a relational database of **policy** and **scholarly** knowledge worldwide



In this project, we analyze data from **1.2 million government policy documents** from **185 countries** extracting the sources they cite to build a relational database of **policy** and **scholarly** knowledge worldwide to answer three sequential questions:

In this project, we analyze data from **1.2 million government policy documents** from **185 countries** extracting the sources they cite to build a relational database of **policy** and **scholarly** knowledge worldwide to answer three sequential questions:

- To what extent do governments rely on **domestic** versus **foreign** sources of evidence?

In this project, we analyze data from **1.2 million government policy documents** from **185 countries** extracting the sources they cite to build a relational database of **policy** and **scholarly** knowledge worldwide to answer three sequential questions:

- To what extent do governments rely on **domestic** versus **foreign** sources of evidence?
- Which countries' **policy** and **scientific outputs** are **most** frequently **cited** and therefore **visible**?

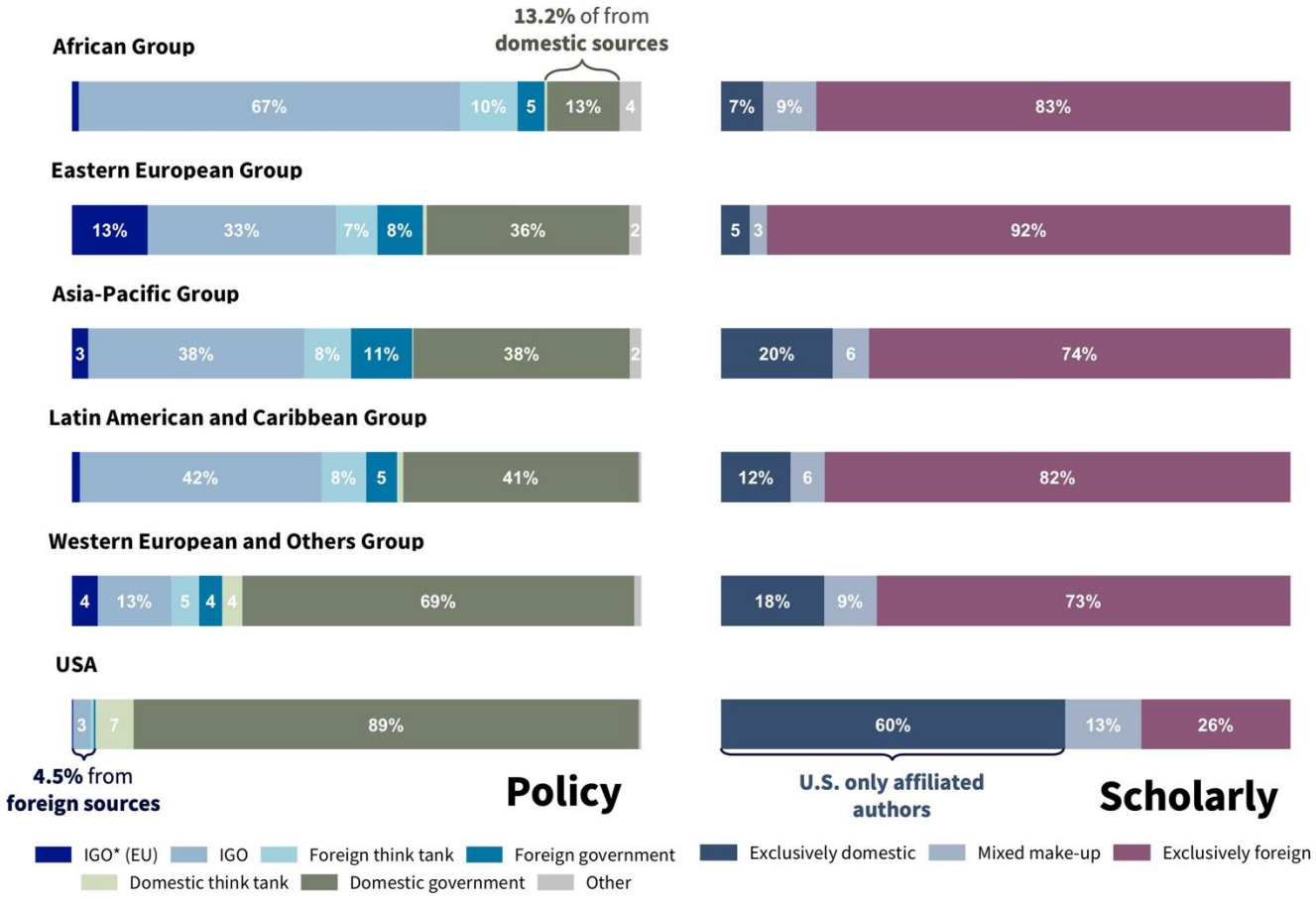
In this project, we analyze data from **1.2 million government policy documents** from **185 countries** extracting the sources they cite to build a relational database of **policy** and **scholarly** knowledge worldwide to answer three sequential questions:

- To what extent do governments rely on **domestic** versus **foreign** sources of evidence?
- Which countries' **policy** and **scientific outputs** are **most** frequently **cited** and therefore **visible**?
- How do citation patterns vary across **policy domains** with **differing knowledge demands**?

What did we find?

1. To what extent do governments rely on **domestic** versus **foreign** sources of evidence?

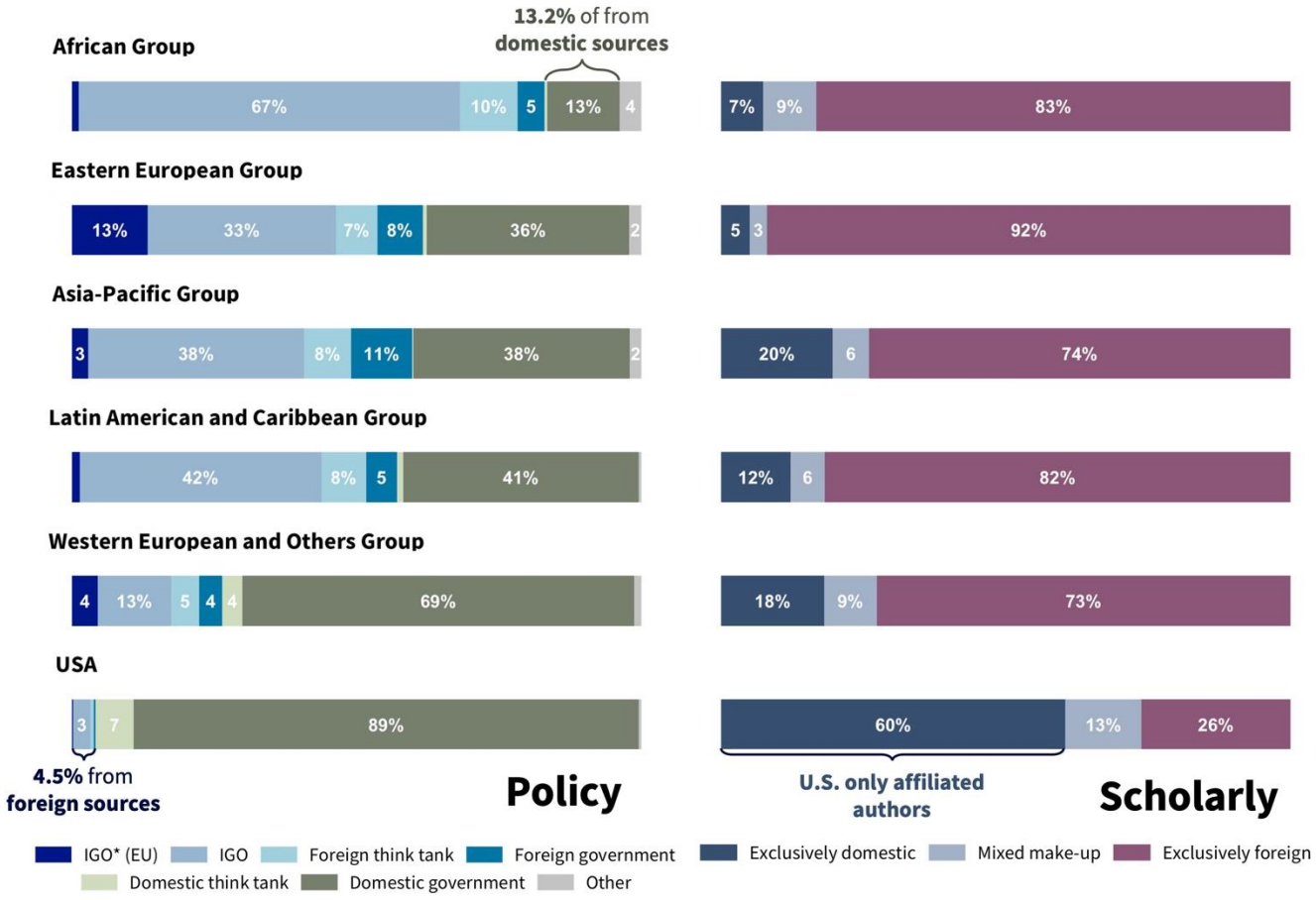
Governments in the global South more frequently cite **foreign** materials



What did we find?

1. To what extent do governments rely on **domestic** versus **foreign** sources of evidence?

Governments in the global South more frequently cite **foreign** materials, while higher-income countries tend to draw from **domestic** expertise

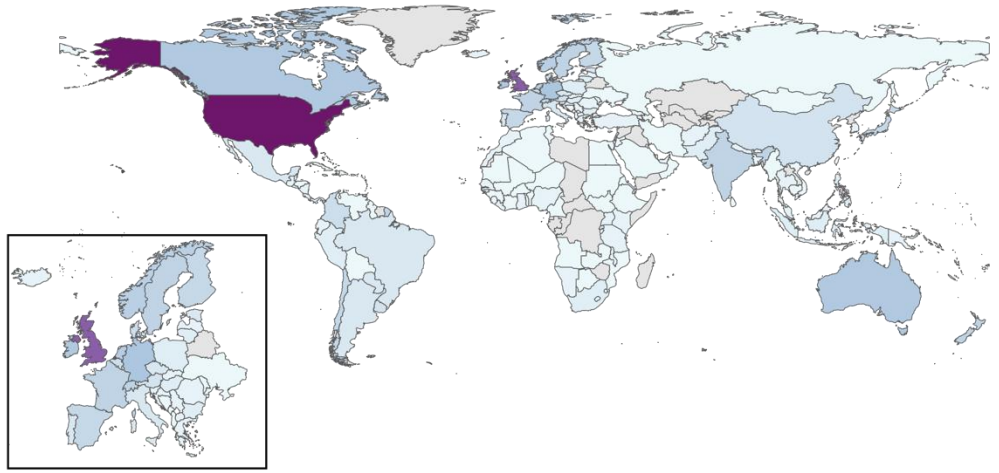


What did we find?

2. Which countries' **policy** and **scientific outputs** are **most** frequently **cited** and therefore **visible**?

Countries in the global North (e.g., US, GB, and EU) lead in **production** of *policy-based* & *academic* ***research*** cited by governments worldwide.

a) Distribution of *H*-indexes government document references



What did we find?

2. Which countries' **policy** and **scientific outputs** are **most** frequently **cited** and therefore **visible**?

When we zoom-into policy 🏛 and academic spaces 📚 :

What did we find?

2. Which countries' **policy** and **scientific outputs** are **most** frequently **cited** and therefore **visible**?

c) Countries with the highest *H*-indexes and share of academic affiliations

COUNTRY	H-I	N. REFS	GROUP
USA	44	10553	WEOG
UK	32	10508	WEOG
Germany	12	986	WEOG
Australia	11	1056	WEOG
Canada	11	1596	WEOG
Ireland	11	978	WEOG
Netherlands	10	663	WEOG
Norway	9	495	WEOG
France	8	706	WEOG
India	8	207	Asia-Pacific Group
New Zealand	8	533	WEOG
Sweden	8	967	WEOG
Switzerland	8	411	WEOG
Chile	7	171	GRULAC
Finland	7	502	WEOG
Japan	7	494	Asia-Pacific Group
Spain	7	628	WEOG
Belgium	6	412	WEOG
Colombia	6	283	GRULAC
Denmark	6	129	WEOG

Government-to-government references

COUNTRY	%. DOCS	N. PAPERS	GROUP
USA	43.4%	1306064	WEOG
UK	13.3%	399890	WEOG
Canada	7.6%	227790	WEOG
Australia	6.6%	199213	WEOG
Germany	6.4%	191658	WEOG
Spain	4.9%	147406	WEOG
France	4.8%	143340	WEOG
Netherlands	4.2%	127181	WEOG
Italy	4%	120763	WEOG
China	3.3%	100288	Asia-Pacific Group
Sweden	3.3%	98589	WEOG
Japan	3.1%	94955	Asia-Pacific Group
Switzerland	2.4%	73578	WEOG
Belgium	1.9%	57785	WEOG
Denmark	1.8%	54042	WEOG
Brazil	1.8%	52669	GRULAC
Finland	1.7%	51935	WEOG
Norway	1.7%	50491	WEOG
India	1.6%	48793	Asia-Pacific Group
New Zealand	1.3%	39300	WEOG

Docs with author based at an insitution in country

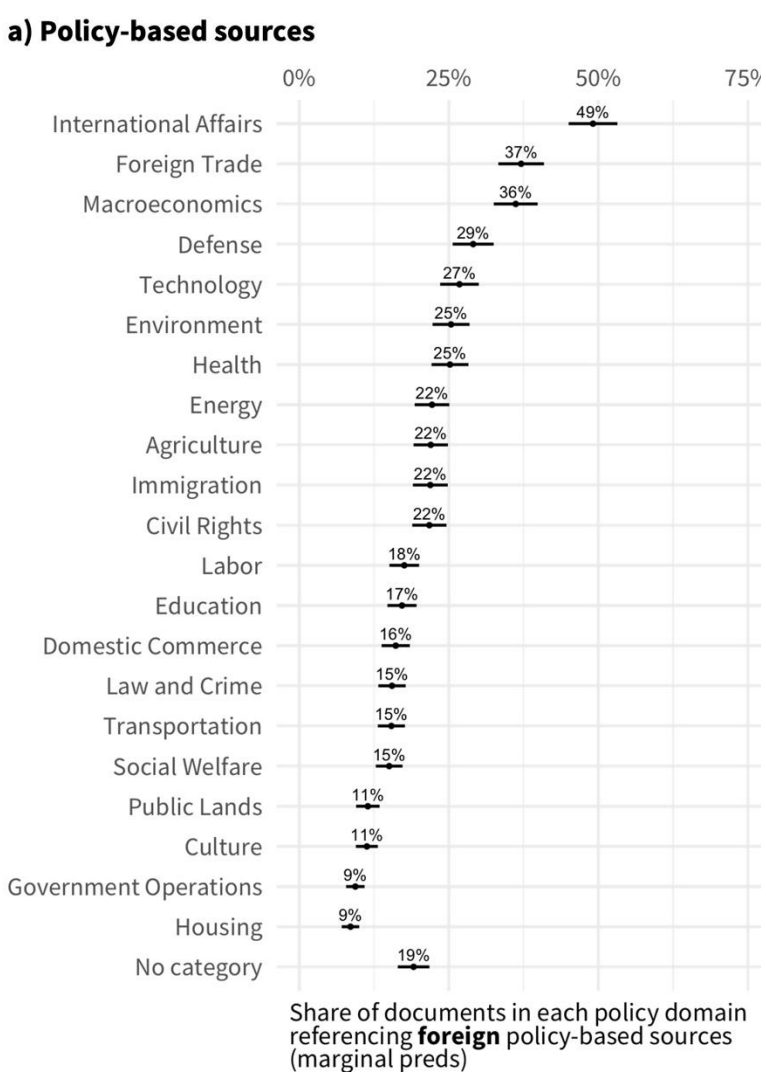
When we zoom-into policy 🏛 and academic spaces 📚 :

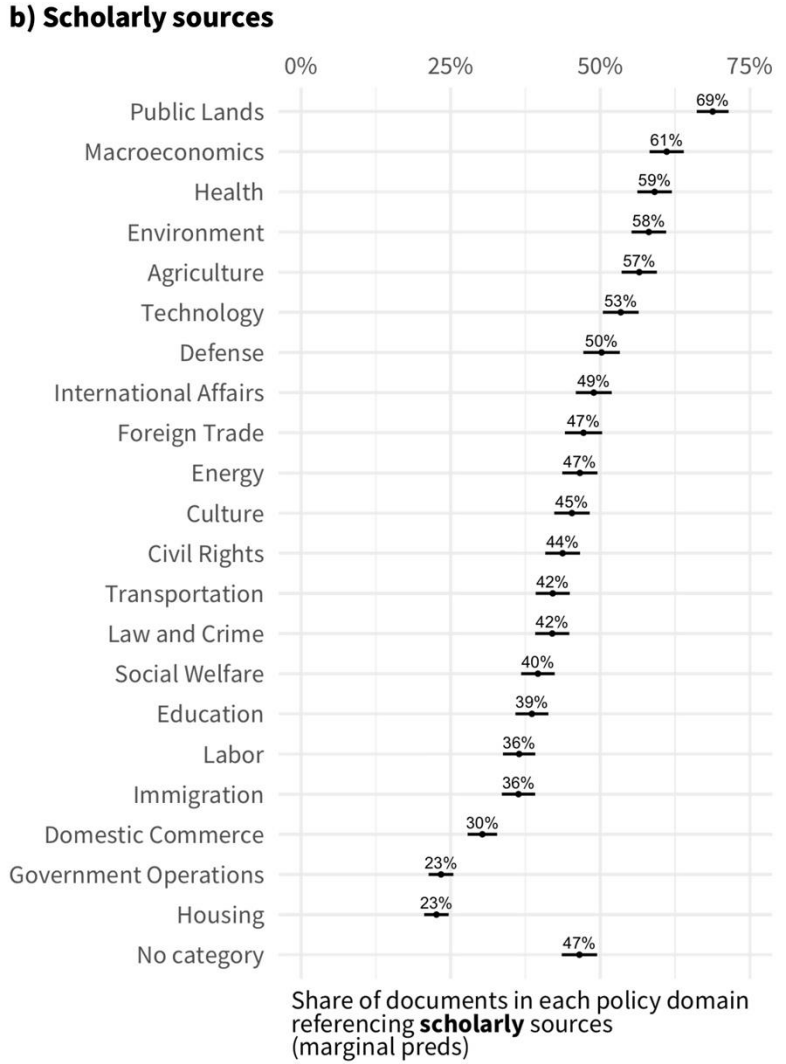
- 17 of the top 20 **countries are shared across** *policy-based* and *scholarly* outputs (largely *WEOG*)

What did we find?

3. How do citation patterns vary across **policy domains** with **differing knowledge demands**?

While **policy domains** have different **evidence orientations**

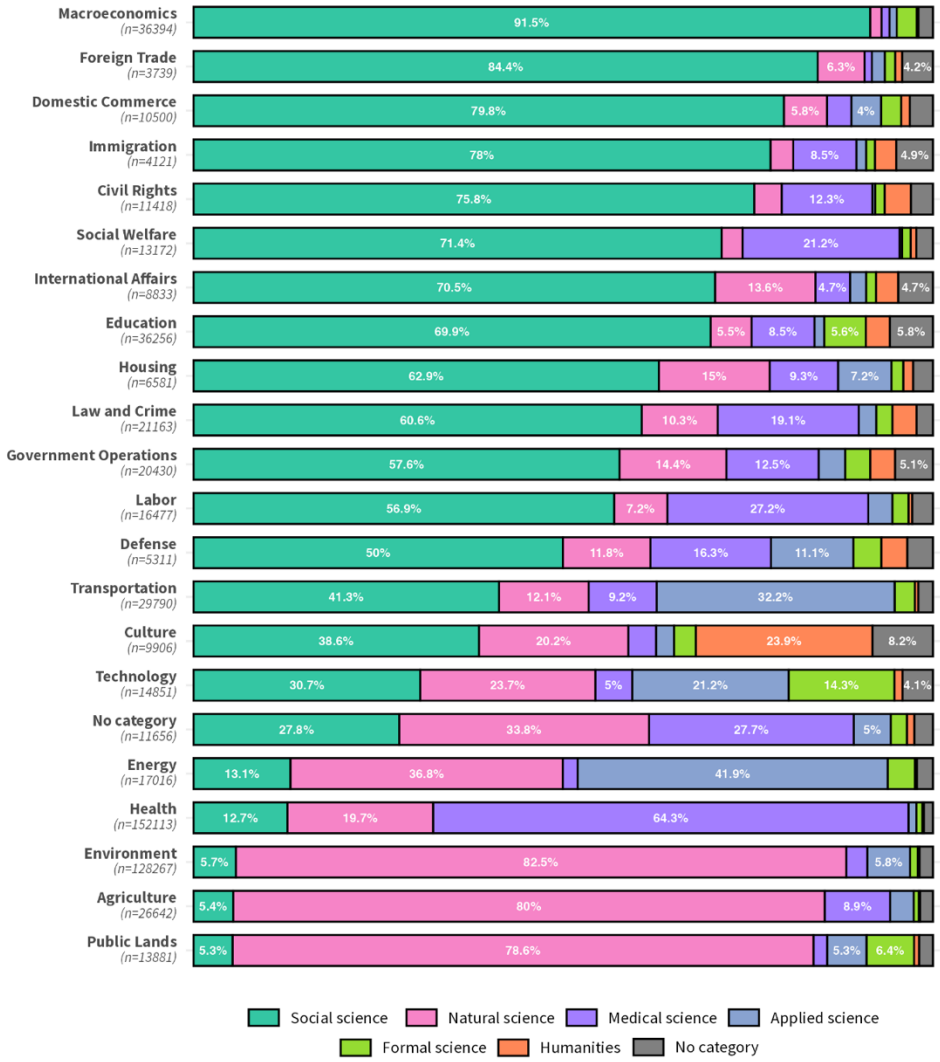




What did we find?

3. How do citation patterns vary across **policy domains** with **differing knowledge demands**?

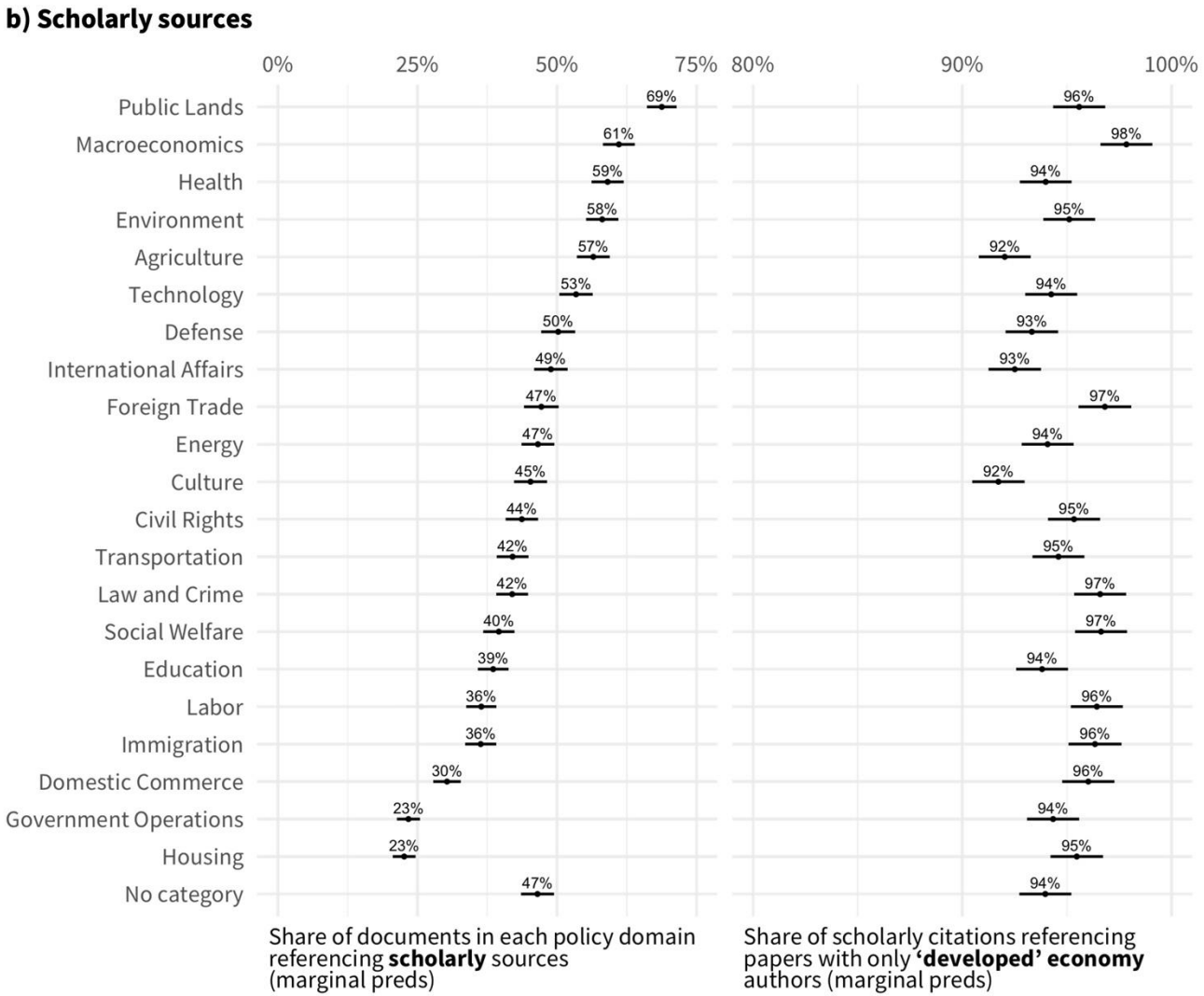
While **policy domains** have different **evidence orientations**



What did we find?

3. How do citation patterns vary across **policy domains** with **differing knowledge demands**?

While **policy domains** have different **evidence orientations**, the pattern holds: evidence from economically powerful nations is consistently more prominent.



institutions

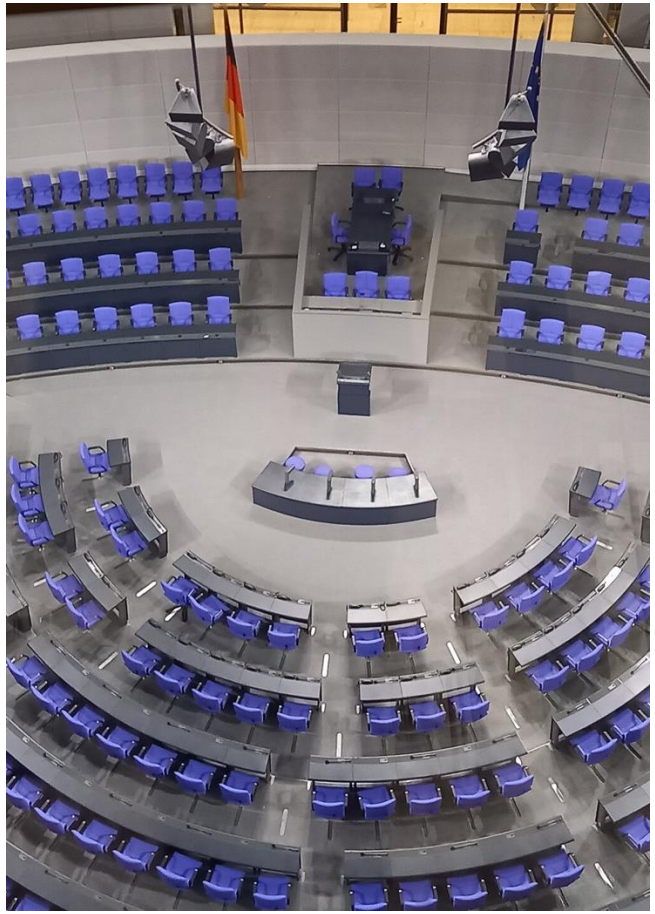
The Bundestag Expert Witness Tracker (BEWIT): A database of German Bundestag public expert hearings

Sebastian Ramirez-Ruiz (Hertie School)*

Abstract. In democratic systems, legislators tackle complex policy challenges while juggling limited time, attention, and information, alongside pressures from constituents and interest groups. Although they frequently rely on external expertise, we still know surprisingly little about who is formally invited to contribute knowledge to the legislative process. This paper introduces the Bundestag Expert Witness Tracker (BEWIT), a novel, hand-curated dataset documenting over 11,000 expert-affiliation pairs from 1,804 public committee hearings in the German Bundestag between 2009 and 2024. The database captures detailed metadata on hearings, expert identities, organizational affiliations, and—where available—linked *Lobbyregister* records and academic researcher profiles. These hearings offer a rare, observable arena to study how legislators curate informational inputs across diverse policy domains. To demonstrate the empirical potential of BEWIT, I present two applications: one that analyzes the disciplinary and institutional composition of academic experts in committee hearings, and another that examines whether a procedural reform in the middle of the 20th legislative period—requiring disclosure of which party invited each expert—correlates with shifts in the makeup of the witness pool. This resource opens new avenues for investigating how democratic institutions filter expertise, navigate competing demands, and structure the informational foundations of policymaking.

Keywords. Public committee hearings; Legislative data; German Bundestag; Evidence-based policymaking; Expert-information; Interest groups; Knowledge utilization

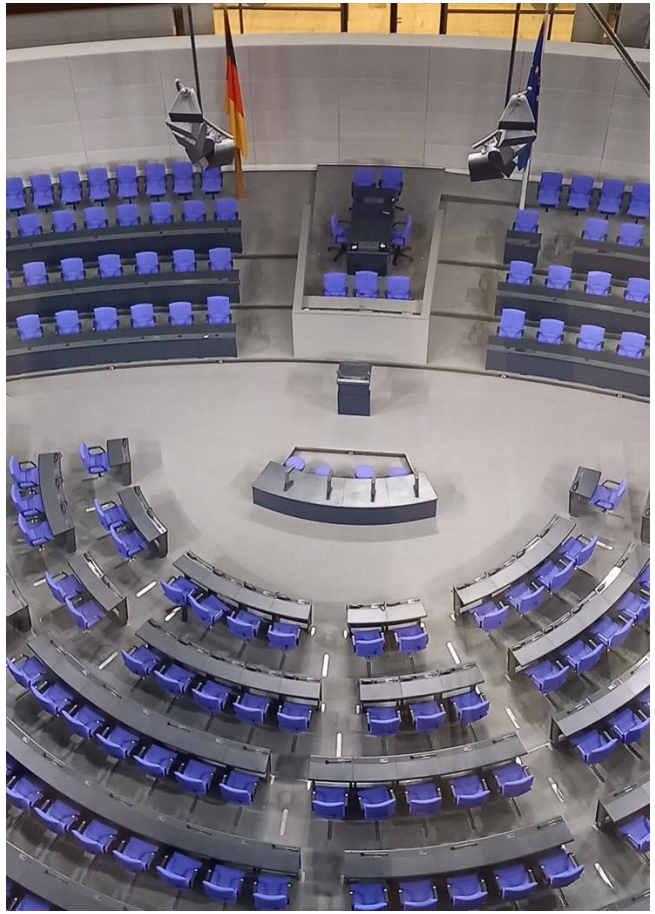
Public expert hearings in the *Bundestag*



- Public hearings serve a crucial function in **supporting legislative information acquisition**, a foundational task for committees.

(Krehbiel, 1992; Strøm, 1998; Baron, 2000)

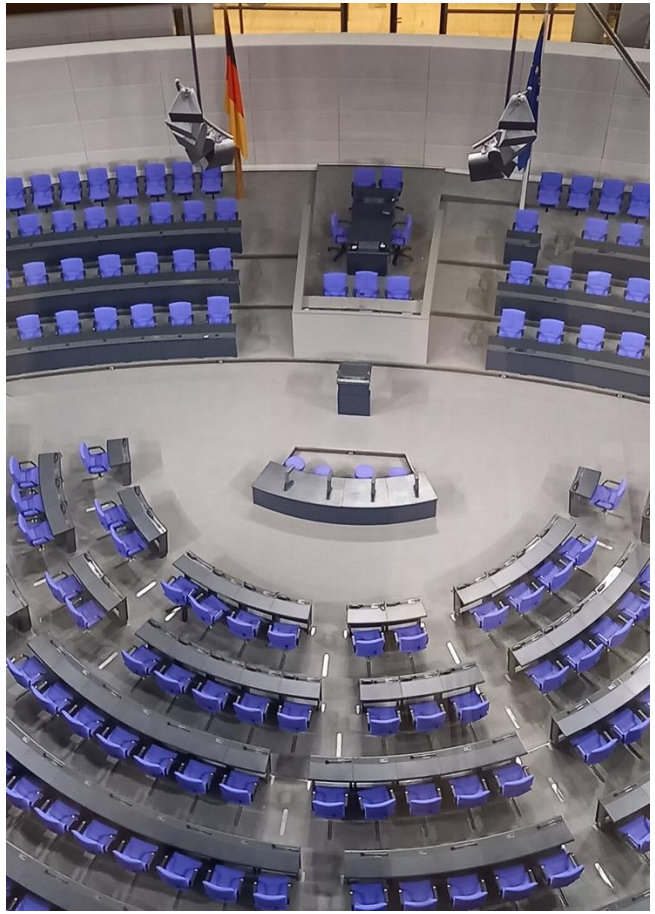
Public expert hearings in the *Bundestag*



- Public hearings serve a crucial function in **supporting legislative information acquisition**, a foundational task for committees.
- Hearings also serve **strategic** and **symbolic** functions (e.g., grandstanding).

(Dhungel and Linhart, 2014 Hünermund, 2020; Geddes, 2024; Park, 2017, 2023)

Public expert hearings in the *Bundestag*

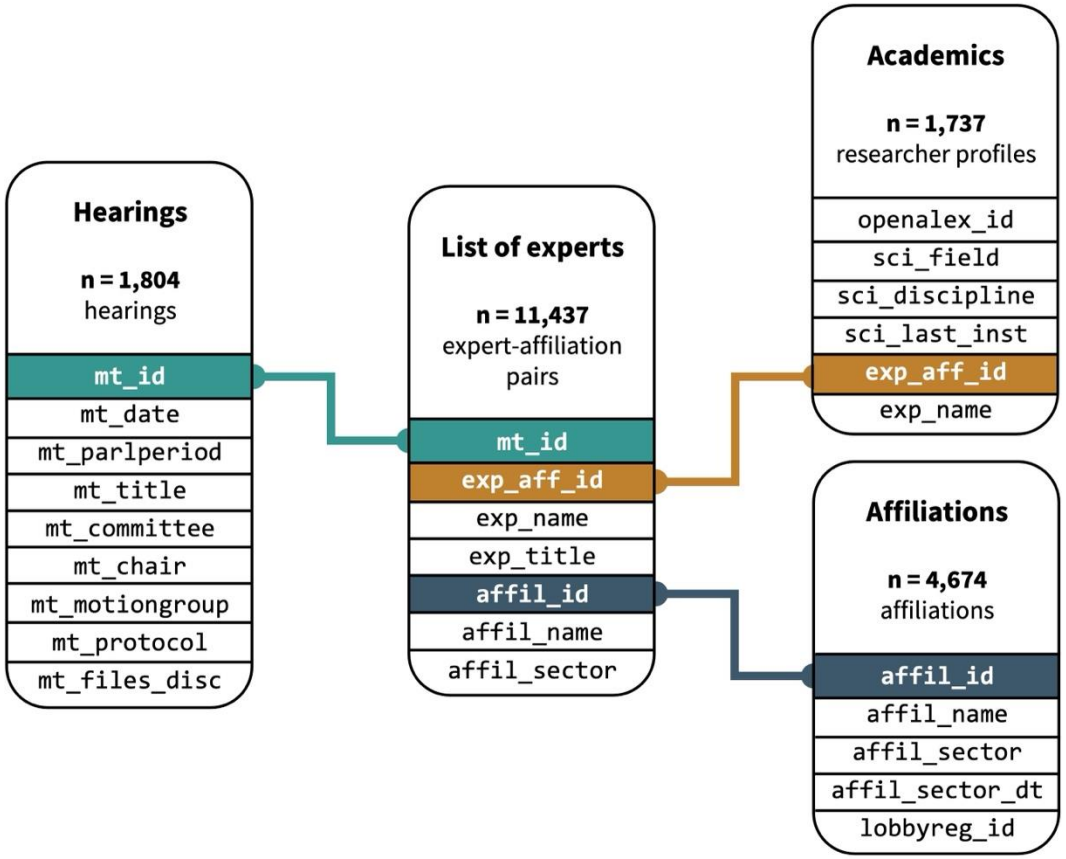


- Public hearings serve a crucial function in **supporting legislative information acquisition**, a foundational task for committees.
- Hearings also serve **strategic** and **symbolic** functions (e.g., grandstanding).
- Unlike the U.S. Congress, where hearings often serve additional purposes such as **oversight** or **confirmations**, Bundestag hearings are more narrowly oriented toward **issue-specific deliberation** and **policy formulation**.

In this project, I **construct** and **present** a hand-curated **database of all public expert committee hearings** in the German Bundestag between 2009 to 2024 and merge it with metadata on the **expert witness** identities (**academic** and **lobbying** profiles)

- **1,804** public expert hearings
- **>18K** expert appearances
- across **4 legislative terms**
- **>11K** expert-affiliation pairs

In this project, I **construct** and **present** a hand-curated **database of all public expert committee hearings** in the German Bundestag between 2009 to 2024 and merge it with metadata on the **expert witness** identities (**academic** and **lobbying** profiles)



In this project, I **construct** and **present** a hand-curated **database of all public expert committee hearings** in the German Bundestag between 2009 to 2024 and merge it with metadata on the **expert witness** identities (**academic** and **lobbying** profiles)

To illustrate the **value** of this resource I answer some questions:

In this project, I **construct** and **present** a hand-curated **database of all public expert committee hearings** in the German Bundestag between 2009 to 2024 and merge it with metadata on the **expert witness** identities (**academic** and **lobbying** profiles)

To illustrate the **value** of this resource I answer some questions:

- **Who** gets **invited** (*platformed*) to testify in parliament by legislators?

In this project, I **construct** and **present** a hand-curated **database of all public expert committee hearings** in the German Bundestag between 2009 to 2024 and merge it with metadata on the **expert witness** identities (**academic** and **lobbying** profiles)

To illustrate the **value** of this resource I answer some questions:

- **Who** gets **invited** (*platformed*) to testify in parliament by legislators?
- What kinds of **knowledge** are seen as **authoritative** across policy areas?

In this project, I **construct** and **present** a hand-curated **database of all public expert committee hearings** in the German Bundestag between 2009 to 2024 and merge it with metadata on the **expert witness** identities (**academic** and **lobbying** profiles)

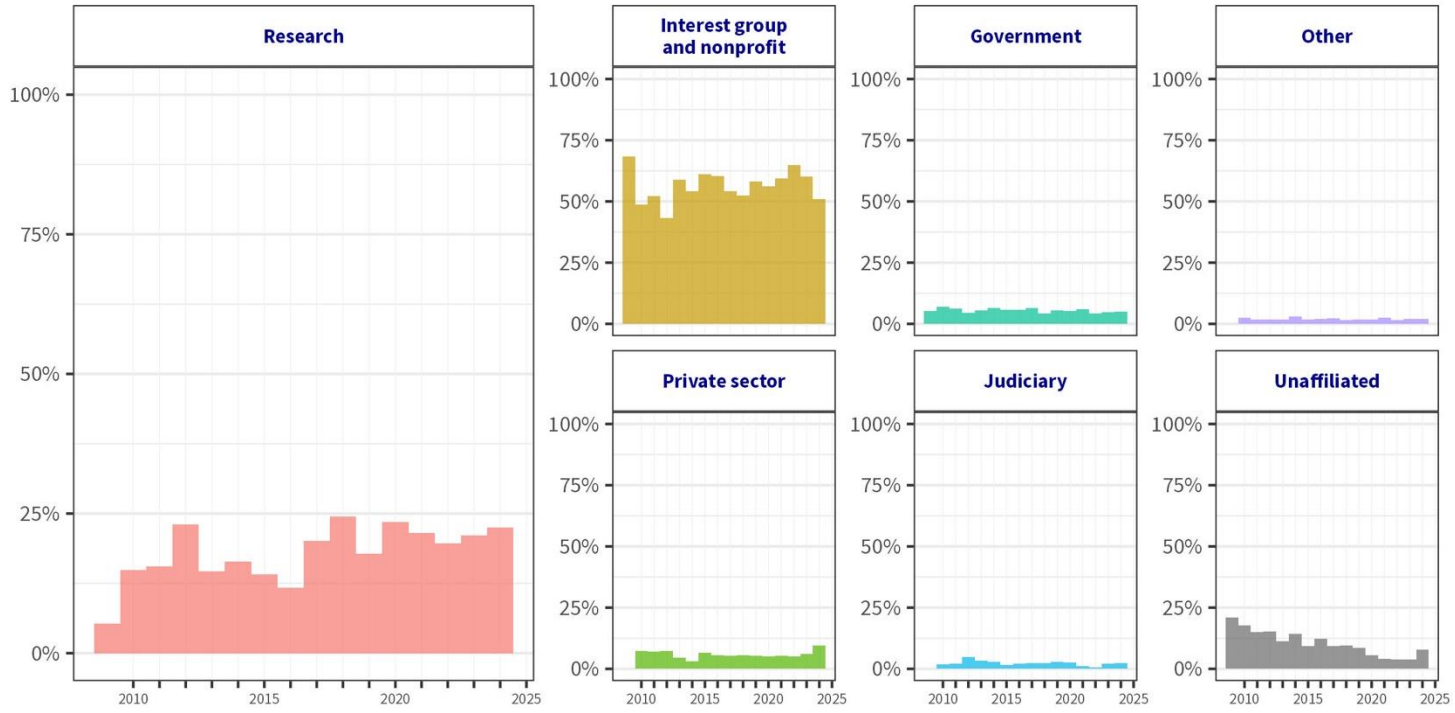
To illustrate the **value** of this resource I answer some questions:

- **Who** gets **invited** (*platformed*) to testify in parliament by legislators?
- What kinds of **knowledge** are seen as **authoritative** across policy areas?
- Do institutional **changes in transparency** affect the **composition** of expert input?

What did I find?

1. **Who** gets **invited** (*platformed*) to testify in parliament by legislators?

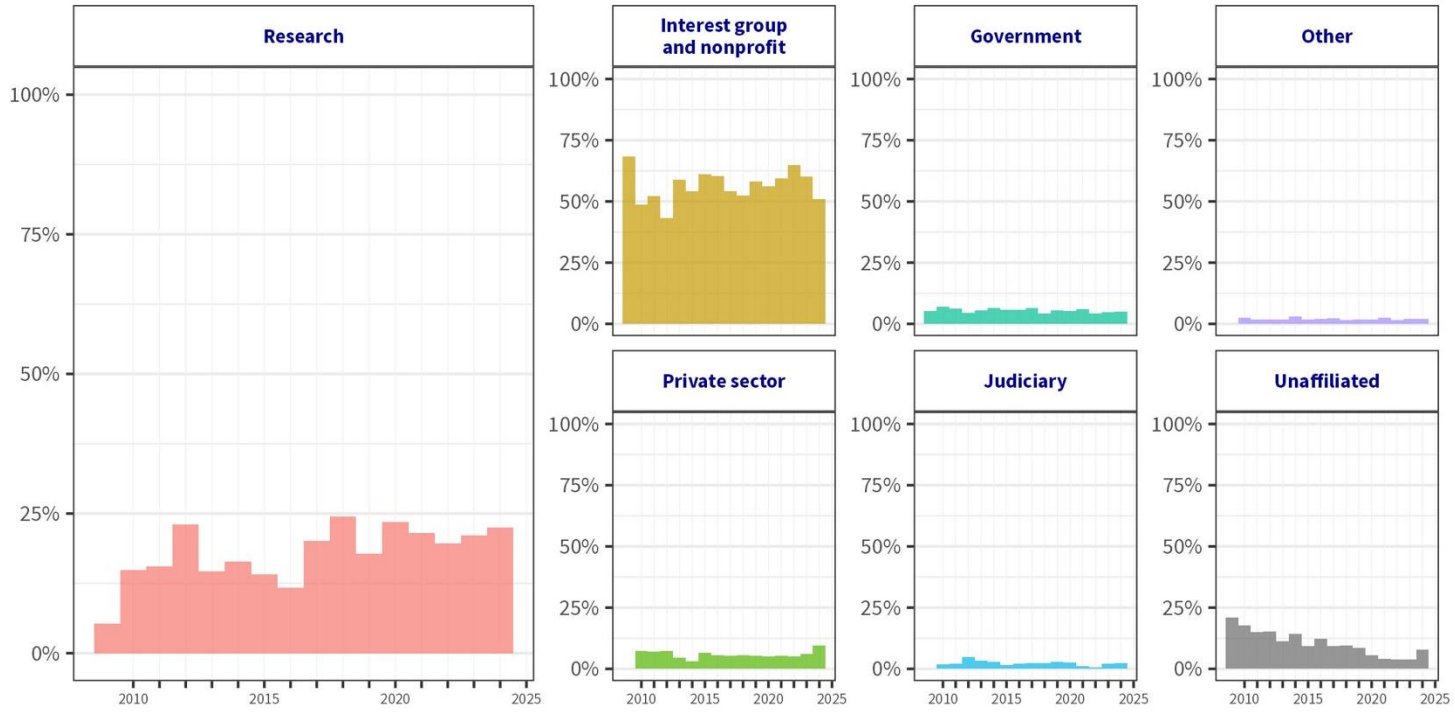
More than half represent **interest groups**



What did I find?

1. **Who** gets **invited** (*platformed*) to testify in parliament by legislators?

Researchers account for **more than 20%** of experts invited to these hearings

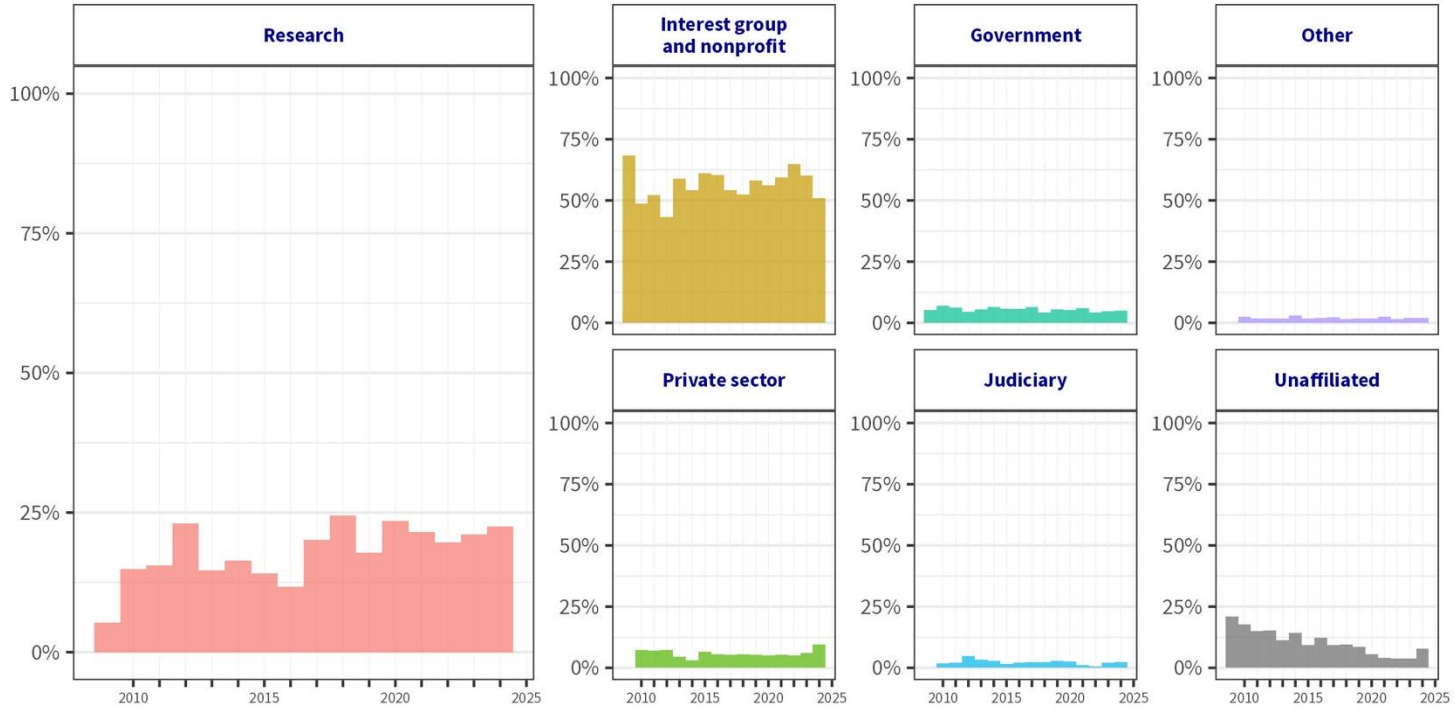


What did I find?

1. **Who** gets **invited** (*platformed*) to testify in parliament by legislators?

Researchers account for **more than 20%** of experts invited to these hearings

*compared to 8% in the U.S. from (Ban, Park, and You, 2023)



What did I find?

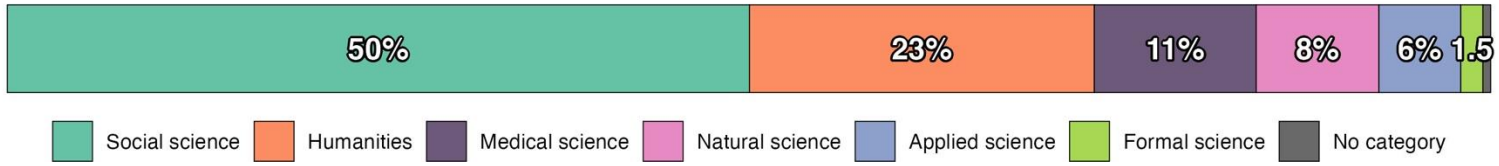
2. What kinds of **knowledge** are seen as **authoritative** across policy areas?

Committees differ in the **extent** to which they draw **academics** and how **interdisciplinary** they draw from

Committee	Experts	Res	Res (%)	Div indx.
European Union Affairs	146	100	68.5%	0.42
Foreign Affairs	51	32	62.7%	0.19
Petitions	7	4	57.1%	0.00
Budget	349	186	53.3%	0.44
Main Committee	40	18	45.0%	0.69
Defence	80	33	41.2%	0.65
Education, Research and Technology Assessment	285	102	35.8%	0.80
Internal Affairs and Community	1201	420	35.0%	0.61
Digital Affairs	205	68	33.2%	0.75
Legal Affairs	2112	595	28.2%	0.57
Human Rights and Humanitarian Aid	265	72	27.2%	0.30
Food and Agriculture	419	106	25.3%	0.63
Finance	1661	411	24.7%	0.44
Economic Cooperation and Development	142	35	24.6%	0.42
Health	1416	342	24.2%	0.68
Family Affairs, Senior Citizens, Women and Youth	536	129	24.1%	0.64
Economic Affairs	884	205	23.2%	0.57
Cultural and Media Affairs	209	42	20.1%	0.49
Env, Nature Cons, Nuclear Safety and Consumer Prot	704	128	18.2%	0.71
Sports	195	29	14.9%	0.44
Transport	741	104	14.0%	0.74
Climate Action and Energy	408	54	13.2%	0.63
Housing, Urban Development, Building and Local Government	268	35	13.1%	0.75
Tourism	258	31	12.0%	0.50
Labor and Social Affairs	1729	174	10.1%	0.66

What did I find?

2. What kinds of **knowledge** are seen as **authoritative** across policy areas?



Social scientists and **legal scholars** are the most common academic expert 'types'

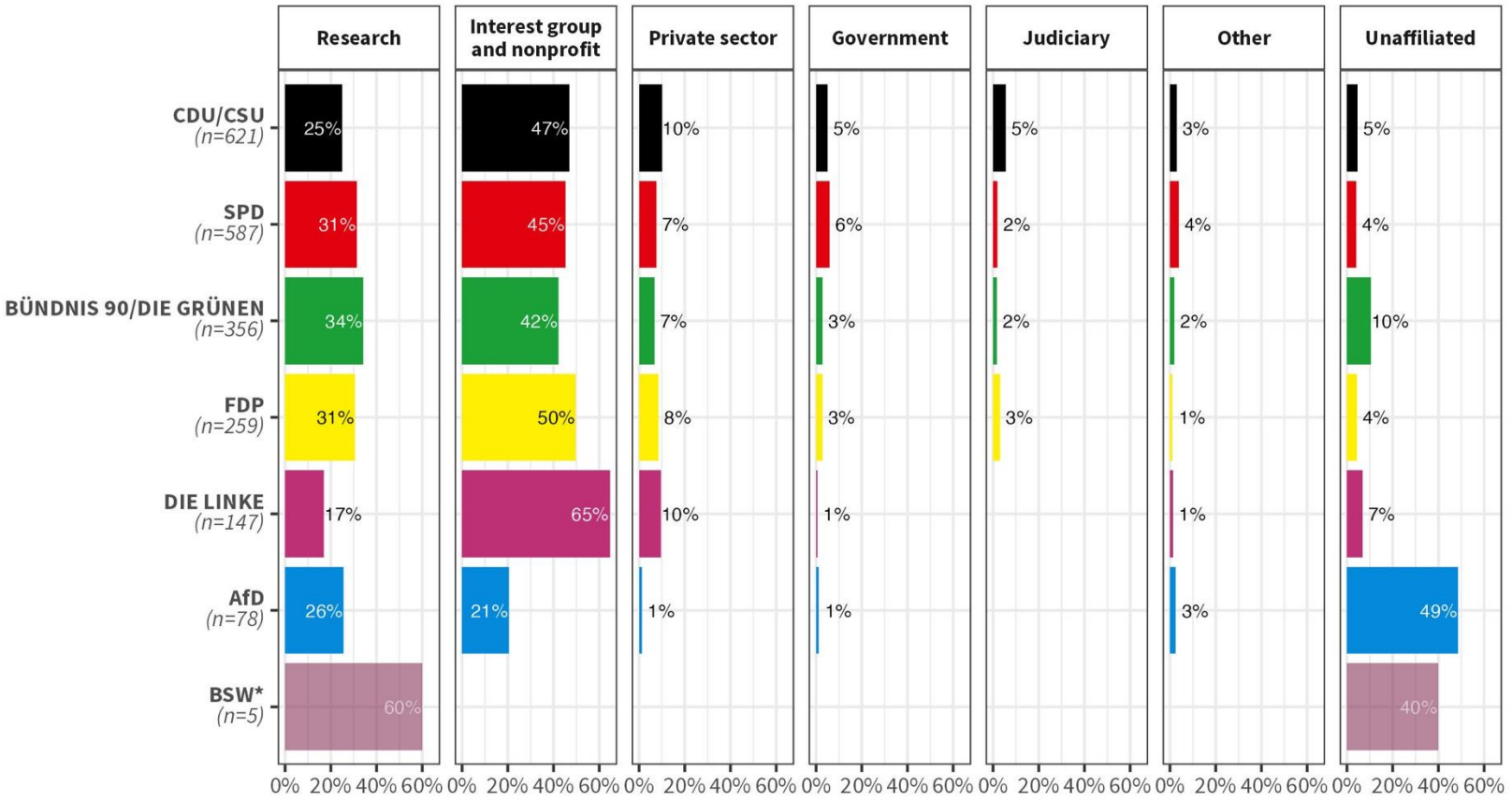
What did I find?

3. Do institutional **changes in transparency** affect the **composition** of expert input?

Starting on **January 1, 2023** (middle of the 20th term), the Bundestag began publishing **which** parliamentary **group invited** each **witness** to testify.

What did I find?

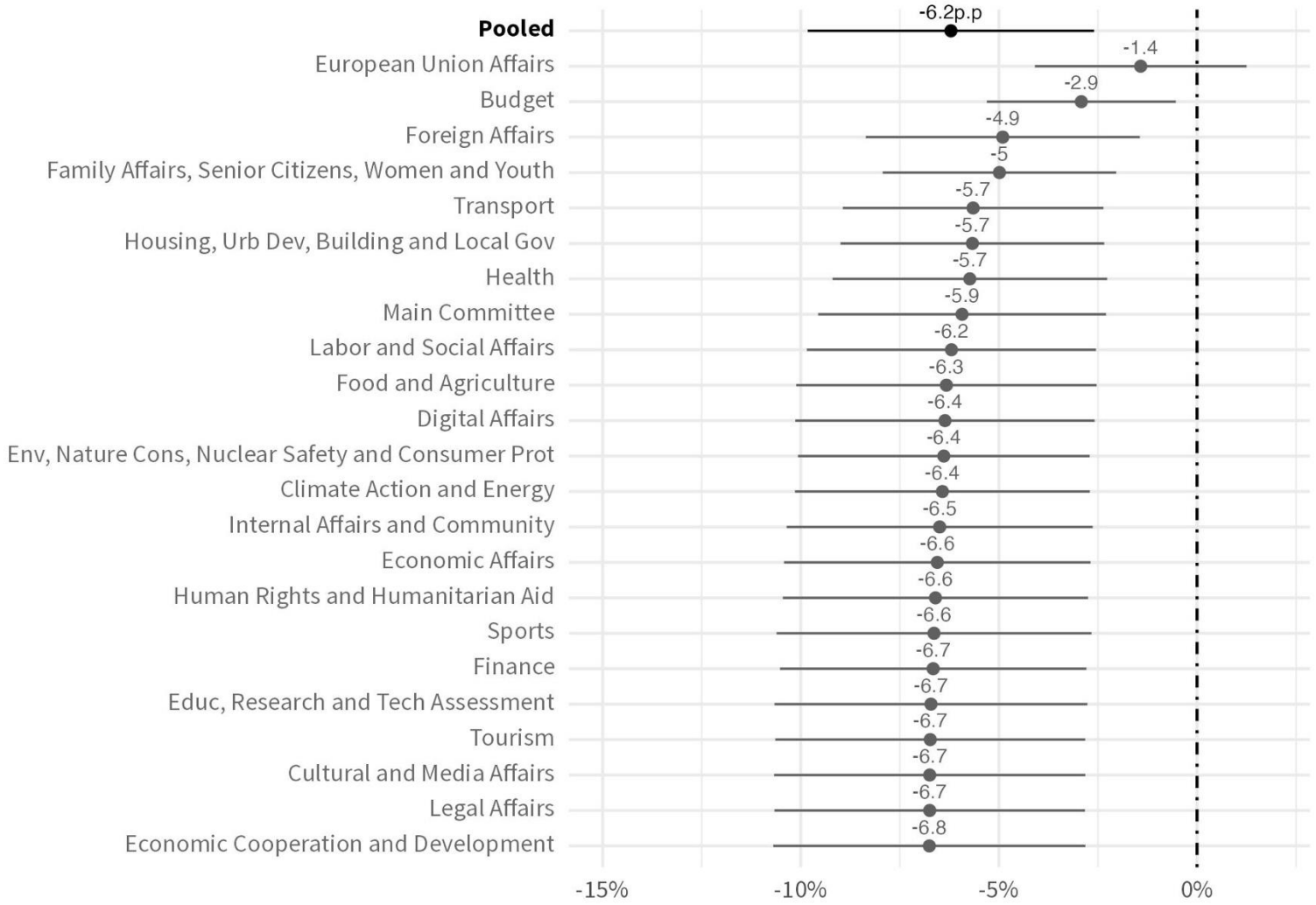
3. Do institutional **changes in transparency** affect the **composition** of expert input?



What did I find?

The change in **protocol** is associated with a measurable **decline** in interest-group witnesses

3. Do institutional **changes in transparency** affect the **composition** of expert input?



competencies

Measuring scientific evidence consumption literacy for public policy: Development and validation of the INSPIRE inventory

Sebastian Ramirez-Ruiz (Hertie School)*
Simon Munzert (Hertie School)

Abstract. Being able to make sense of scientific evidence is a critical component in shaping effective public policy, enabling both policymakers and the public to make informed decisions on complex issues. However, existing measures of this skillset may not fully address the specific methodological and applied demands of public policy contexts. We develop the *Inventory for Numeracy, Statistics, and Policy-oriented Inference and REasoning* (INSPIRE), a comprehensive measurement instrument to assess competence in scientific evidence consumption in relation to public policy. The instrument integrates knowledge of statistical reasoning, data and visualization literacy, causal reasoning and inference, and the scientific method, alongside the ability to critically evaluate scientific information pertinent to policy debates. Using Item Response Theory (IRT), we systematically assess an initial item pool, resulting in an inventory of 30 items. We assessed the inventory's psychometric and substantive validity across three samples: a general population, policy students and professionals in a data science training program, and participants in a pre-election forecasting study. This allowed us to examine its performance across different groups and contexts. Results indicate good internal consistency and evidence of construct, criterion, and predictive validity. We close with considerations for potential use of the item bank in research and applied settings.

Keywords. scientific evidence consumption | scientific reasoning | public policy | evidence-informed policymaking | item response theory (IRT) | Test instrument

In this project, we **develop**, **validate**, and **introduce** an inventory designed to assess essential **competencies** in **scientific evidence consumption** within the policy domain.

In this project, we **develop**, **validate**, and **introduce** an inventory designed to assess essential **competencies** in **scientific evidence consumption** within the policy domain.

- Despite **recognized importance** of these skills, we are **ill equipped** to **measure** them in policy contexts. Existing instruments are:

(see Opitz, Heene, and Fischer, (2017) for a review of instruments)

In this project, we **develop**, **validate**, and **introduce** an inventory designed to assess essential **competencies** in **scientific evidence consumption** within the policy domain.

- Despite **recognized importance** of these skills, we are ill equipped to **measure** them in policy contexts. Existing instruments are:
  - designed for educational settings, particularly **K-12 students**

(see Opitz, Heene, and Fischer, (2017) for a review of instruments)

In this project, we **develop**, **validate**, and **introduce** an inventory designed to assess essential **competencies** in **scientific evidence consumption** within the policy domain.

- Despite **recognized importance** of these skills, we are ill equipped to **measure** them in policy contexts. Existing instruments are:
  - designed for educational settings, particularly **K-12 students**
  - focus on **natural** and **applied sciences**

(see Opitz, Heene, and Fischer, (2017) for a review of instruments)

In this project, we **develop**, **validate**, and **introduce** an inventory designed to assess essential **competencies** in **scientific evidence consumption** within the policy domain.

- Despite **recognized importance** of these skills, we are ill equipped to **measure** them in policy contexts. Existing instruments are:
  - designed for educational settings, particularly **K-12 students**
  - focus on **natural** and **applied sciences**
  - concerned with individuals acting as **basic researchers** not the role of **scientifically knowledgeable practitioners**

(see Opitz, Heene, and Fischer, (2017) for a review of instruments)

What did we do?

We **compiled** and **drafted** an **initial pool** of 52 items across **four conceptual components**.

What did we do?

We **compiled** and **drafted** an **initial pool** of 52 items across **four conceptual components**.

*a simplified definition of what we are trying to get to

What did we do?

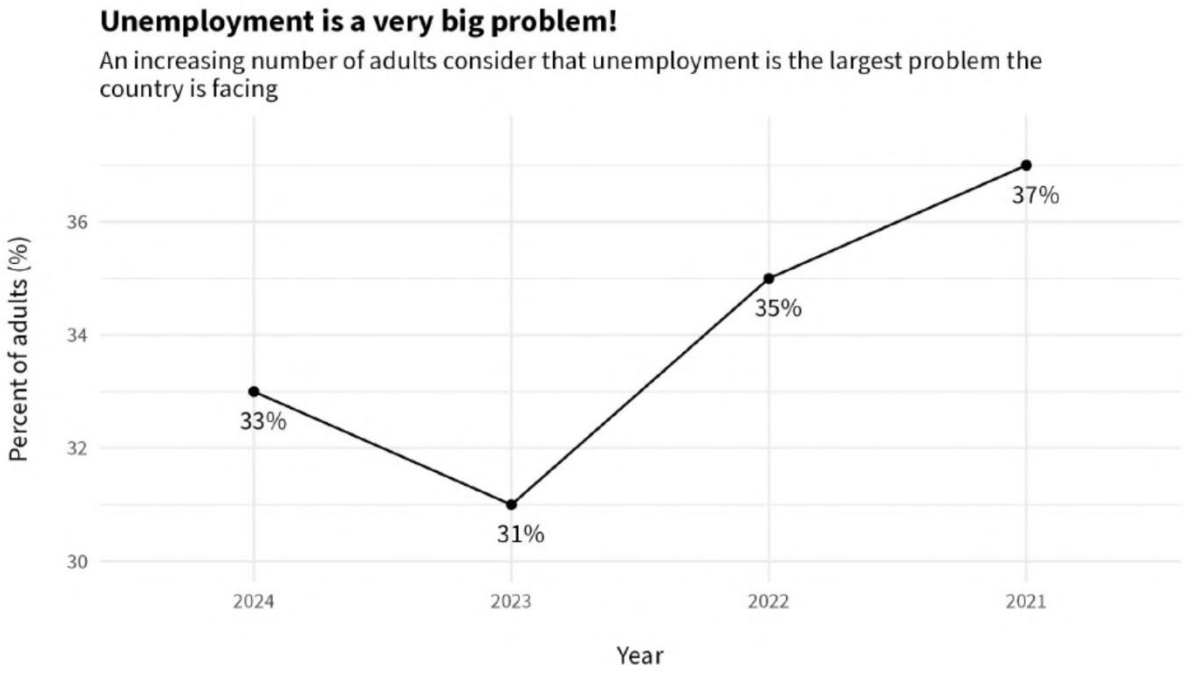
We **compiled** and **drafted** an **initial pool** of 52 items across **four conceptual components**.

Scientific evidence consumption literacy as the **ability** to make sense of **scientific** research and **data** in everyday life.

*a simplified definition of what we are trying to get to

DVL7. Flawed Visuals 1

Examine the graph carefully. Which IS NOT a problematic feature?



A. The x-axis is inverted
B. **Data points shown with low precision ✓**
C. Axis does not start at 0%
D. Subtitle is inconsistent with trend

CR6. Causal Effect of a Policy

Which option best describes a situation where a policy has a causal effect on unemployment?

A. Unemployment fell after the policy was implemented
B. The policy is correlated with changes in the unemployment rate
C. **The policy directly causes changes in the unemployment rate, beyond what would have happened without the policy ✓**
D. The policy was implemented at the same time as changes in the unemployment rate

What did we do?

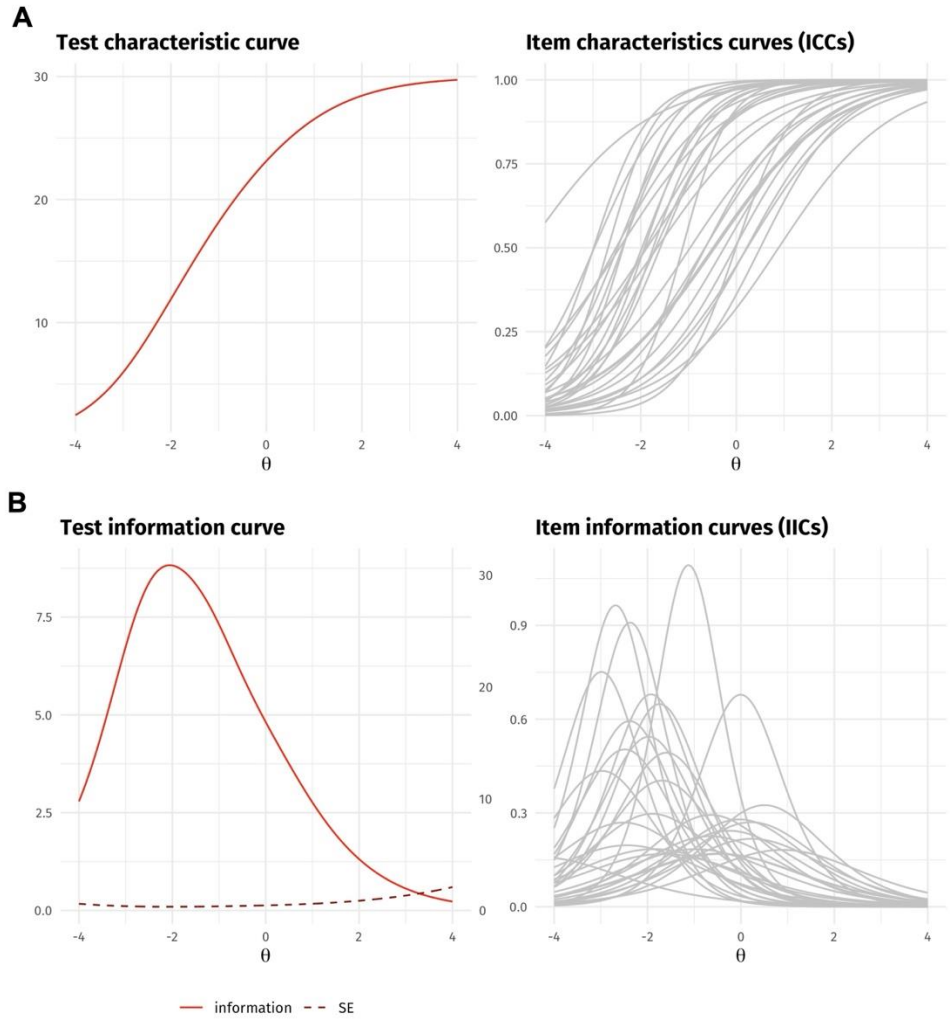
We delivered them in an **online survey** context

470 respondents via Prolific:

- 338 general population participants
- 132 government workers

What did we do?

We delivered them in an **online survey** context, tested the psychometric properties of items (**IRT**)



What did we do?

We delivered them in an **online survey** context, tested the psychometric properties of items (**IRT**), and extracted a **reduced** validated **pool**

Domain and sub-domain	Item code	Item label	2PL Disc (a)	2PL Diff (b)	Percent correct	Median time (s)	N
Scientific Literacy							
Basic science	SEL04	Research credibility	1.73	-2.99	98.0%	14	175
Basic science	SEL03	Quality of evidence	1.91	-2.37	95.0%	18	170
Source reliability	SEL07	Trustworthy data	1.54	-2.39	94.0%	13	190
Source reliability	SEL08	Trustworthy sources	1.42	-2.49	94.0%	24	172
Scientific practice	SEL06	Scientific practice 2	1.65	-1.92	90.0%	21	176
Basic science	SEL01	Scientific hypothesis	1.61	-1.74	88.0%	10	189
Scientific practice	SEL05	Scientific practice 1	0.89	-2.44	85.0%	22	175
Basic science	SEL02	Scientific consensus	1.06	-0.03	50.0%	9	176
Statistical Literacy							
Machine learning models	SML05	ML and perpetuated bias	1.96	-2.68	97.0%	8	220
Machine learning models	SML07	ML application	1.04	-2.52	91.0%	10	210
Base rate	SML03	Base rate 1	1.40	-1.59	88.0%	39	160
Machine learning models	SML04	ML accuracy	1.27	-1.69	84.0%	9	234
Center and spread	SML01	Center and spread	0.82	-1.66	74.0%	36	163
Center and spread	SML02	Sample sizes and uncertainty	0.95	-0.66	65.0%	36	162
Machine learning models	SML06	ML and flawed predictions	0.82	-0.45	60.0%	29	193
Data Literacy							
Visual information	DVL06	Visual information 2	1.32	-2.97	95.0%	24	261
Numeracy	DVL01	Complementary probability 1	1.47	-1.97	90.0%	11	174
Visual information	DVL05	Visual information 1	0.86	-1.87	81.0%	48	241
Visual representation	DVL04	Visual data representation 1	0.99	-0.25	57.0%	33	235
Numeracy	DVL02	Complementary probability 2	1.65	-0.01	50.0%	31	182
Numeracy	DVL03	Percentages	0.93	0.23	48.0%	42	181
Problematic visuals	DVL07	Flawed visuals 1	1.14	0.50	41.0%	58	253
Problematic visuals	DVL08	Flawed visuals 2	0.85	0.89	37.0%	54	245
Causal Reasoning							
Experiments	CR04	Randomized control trials	0.80	-4.38	96.0%	10	160
Causal reasoning	CR05	Confounder	1.09	-1.90	85.0%	24	196
Valid causal conclusions	CR07	Study conclusions 1	2.09	-1.12	82.0%	26	181
Correlation	CR03	Learning from correlation	0.86	-0.97	69.0%	27	192
Causal policy	CR06	Causal effect of a policy	1.08	-0.65	64.0%	27	193
Correlation	CR02	Correlation	0.82	-0.49	60.0%	20	197
Correlation	CR01	Correlation vs. Causation	1.04	0.22	46.0%	10	168

- **discrimination** (a)
  - items that more effectively distinguish between respondents at different levels of latent ability
- **local dependence** (*Yen's* q^3)
  - avoid shared variance not related to trait (e.g., from similar content or item stems/scenarios)

What did we do?

We validated our **measure** with different benchmarks **across** varied **samples**

- Prolific (n=**470**)
- Policy professionals and students (n=**21**)
- Pre-election survey (n=**15k**)

What did we do?

We validated our **measure** with different benchmarks **across** varied **samples**

Scores on the latent trait are correlated to self-reported **familiarity**, **knowledge**, and **ability**

(*Prolific sample*)
Correlation with self-reported knowledge in the following areas:

	Math (arithmetic, linear algebra, calculus)	Probability and statistics	Data analysis	Scientific methods	Cause-and-effect reasoning	Consumption of scientific evidence
Full scale (θ)	0.16	0.21	0.19	0.28	0.26	0.28

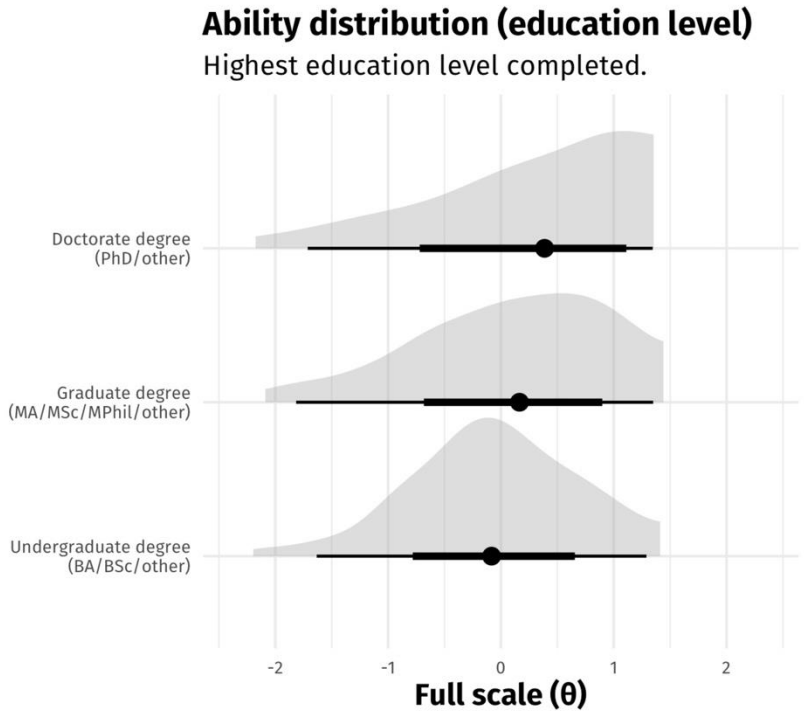
(*Policy student and professional sample*)
Correlation with self-assessment regarding their ability to:

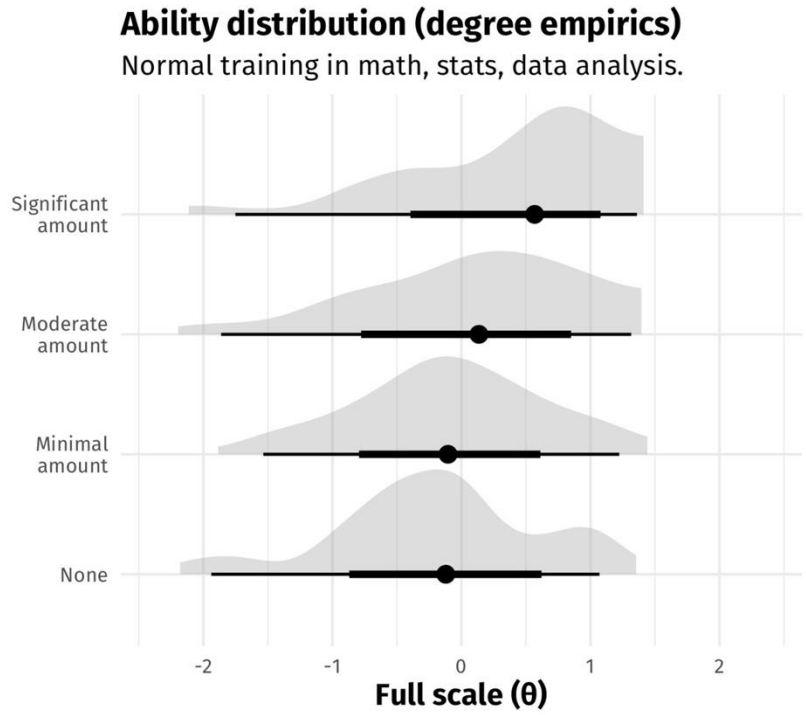
	describe data necessary to answer policy questions	choose methods to answer policy questions	use graphs or math representatic to describe data	convert raw data points and present them	understand technical language in scientific studies	come up with arguments on the basis of evidence	identify research designs more robust for specific questions	design studies to answer policy questions
Full scale (θ)	0.61	0.61	0.4	0.45	0.53	0.34	0.37	0.25

What did we do?

We validated our **measure** with different benchmarks **across** varied **samples**

Scores on the latent trait are correlated to self-reported **familiarity**, **knowledge**, and **ability**

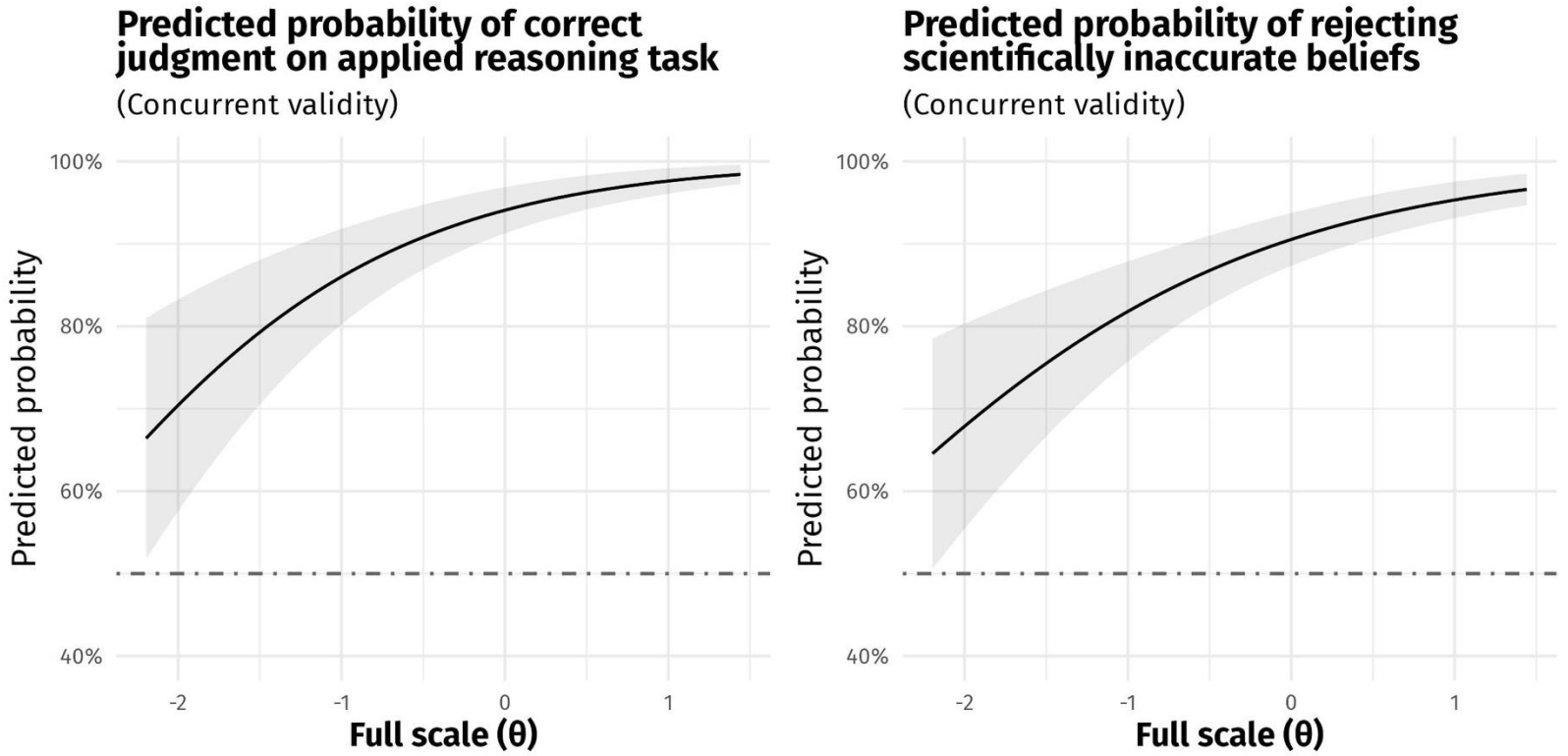




What did we do?

We validated our **measure** with different benchmarks **across** varied **samples**

Scores are strong predictors of individuals' ability to correctly **select** suitable **evidence** for policy questions and **reject** scientifically **inaccurate** beliefs



As I said in my opening, in the dissertation I take a **step back** and ask:

As I said in my opening, in the dissertation I take a **step back** and ask:

"**Would** she **encounter** researchers and be **confronted** with **X** research **finding** in the first place?"

It might depend on structural **inequalities**, **institutional practices**, and some of her **individual characteristics**

(e.g., ideological inclination, consumption capacity and competence)

What counts as **evidence**?

What counts as **evidence**?

It can take **different forms** and come from **different providers**

What counts as **evidence**?

Whose expertise permeates decisions?

What counts as **evidence**?

Whose expertise permeates decisions?

Stories of **power**, **infrastructure**, **ideology**, and scientific **disciplines**

What counts as **evidence**?

Whose expertise permeates decisions?

How does knowledge flow—*if at all*—from research and expert communities to elite decision-makers?

What counts as **evidence**?

Whose expertise permeates decisions?

How does knowledge flow—*if at all*—
from research and expert communities to
elite decision-makers?

So many **different channels**

Informal (e.g., online), artifacts of policymaking, and institutional settings

And **do policymakers have** the **competencies** to work with evidence in the first place?

I do not know, but we can try to **measure**, **monitor**, and **intervene**

And **do policymakers have** the **competencies** to work with evidence in the first place?

Where do we go
from here?

vp and whitehouse

1,641 232

Liked by **bendel_sara** and **others**

vp Today, Vice President Vance and Prime Minister Viktor Orbán of Hungary met at the Vice Presidential Residence.

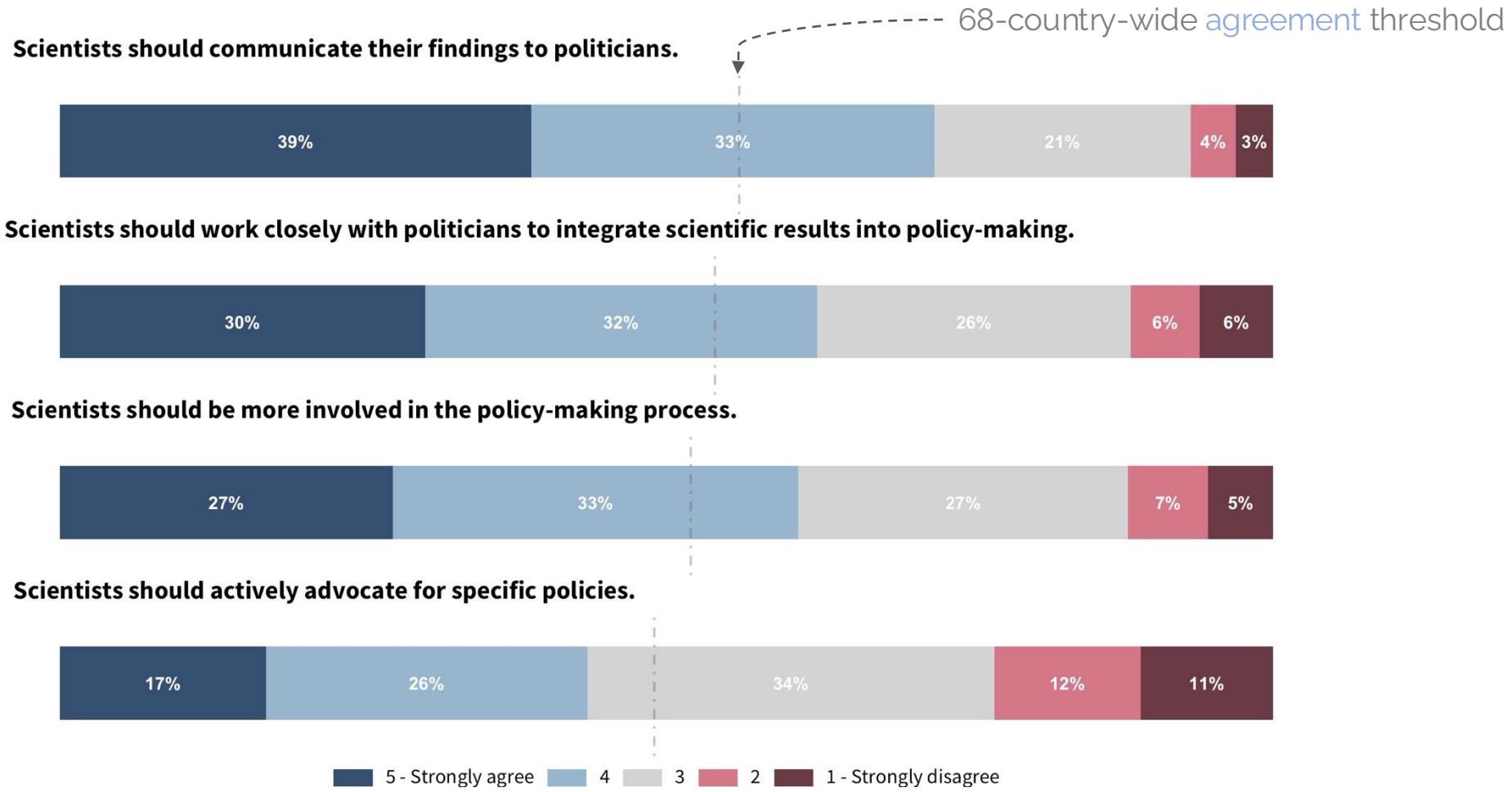
"We have to honestly and **aggressively attack the universities** in this country

...

The universities **do not pursue knowledge** and **truth**, they pursue **deceit** and **lies**"

J.D. Vance
'Universities are the enemy'
Keynote at NatCon 2021

German public



*based on Trust in Science and Science-Related Populism (*TISP*) mega-study data (Mede et al., 2023)

What counts as **evidence**?

Whose expertise permeates decisions?

How does knowledge flow—*if at all*—from research and expert communities to elite decision-makers?

And **do policymakers have** the **competencies** to work with evidence in the first place?

What knowledge gets to be produced?

Data **collection** and **curation**

Data **collection** and **curation**

Research **infrastructure** building

Data **collection** and **curation**

Research **infrastructure** building

Quantitative **description**

Data **collection** and **curation**

Research **infrastructure** building

Quantitative **description**

are **worthwhile scientific endeavors**

Thank you!


seramirezruiz.github.io


Sebastian.ramirezruiz@eui.eu


seramirezruiz.bsky.social

The Evidence Interface

How policymakers encounter, engage with, and make sense of scientific knowledge

Sebastian Ramirez-Ruiz

(Data Science Lab — Hertie School)

Dissertation Defense

November 11, 2025

Hertie School