



MAX-PLANCK-INSTITUT  
FÜR DEMOGRAFISCHE  
FORSCHUNG

MAX PLANCK INSTITUTE  
FOR DEMOGRAPHIC  
RESEARCH

# North and South: Naming practices and the hidden dimension of global disparities in knowledge production

Andrés F. Castro Torres\* & **Diego Alburez-Gutierrez\*\***

\* Center for Demographic Studies (CED), Barcelona

\*\*Max Planck Institute for Demographic Research (MPIDR), Germany

Covenant University, Nigeria

June 21, 2022 - Summer Institutes in Computational Social Science



# Contents

1. Motivation
2. Data and methods
3. Main results
4. Additional results
5. Conclusion and implications



# Introduction

Think about the last article, manuscript, working paper, you worked on...



# Introduction

Think about the last article, manuscript, working paper, you worked on...

1. Did it include empirical data?

Yes -> pass to (2)

**No** -> Thank for your participation



# Introduction

Think about the last article, manuscript, working paper, you worked on...

1. Did it include empirical data?

Yes -> pass to (2)

**No** -> Thank for your participation

2. Did you write the name of the country, region, population under study in the title?



# Introduction

Think about the last article, manuscript, working paper, you worked on...

1. Did it include empirical data?

Yes -> pass to (2)

**No** -> Thank for your participation

2. Did you write the name of the country, region, population under study in the title?

Yes -> raise your hand

# Naming Practices

Original Articles

## Demographic Transition Theory

Dudley Kirk

Pages 361-387 | Published online: 04 Jun 2010

## Modelling the proximate determinants of fertility for Brazil: the advent of competing preferences

Raquel Zanatta Coutinho\*  
Andre Braz Golgher\*\*

Original Article | Published: 01 September 2018


## Armed Conflict and Fertility in Colombia, 2000–2010

[Andrés Felipe Castro Torres](#)  & [B. Piedad Urdinola](#)

*Population Research and Policy Review* **38**, 173–213(2019) | [Cite this article](#)

## The proximate determinants of fertility

John Bongaarts \*

[Show more](#) 

[https://doi.org/10.1016/0160-791X\(87\)90003-0](https://doi.org/10.1016/0160-791X(87)90003-0)





[Get rights and content](#)

## Age patterns of mortality for older women: An analysis using the age-specific rate of mortality change with age

Shiro Horiuchi & Ansley J. Coale

### RESEARCH ARTICLE

## Tracking the reach of COVID-19 kin loss with a bereavement multiplier applied to the United States

 Ashton M. Verdery,  Emily Smith-Greenaway,  Rachel Margolis, and  Jonathan Daw



# Naming Practices

Original Articles

## Demographic Transition Theory

Dudley Kirk

Pages 361-387 | Published online: 04 Jun 2010

## Modelling the proximate determinants of fertility for Brazil: the advent of competing preferences

Raquel Zanatta Coutinho\*  
Andre Braz Golgher\*\*

Original Article | Published: 01 September 2018

## Armed Conflict and Fertility in Colombia, 2000–2010

[Andrés Felipe Castro Torres](#) & [B. Piedad Urdinola](#)

*Population Research and Policy Review* **38**, 173–213(2019) | [Cite this article](#)

## The proximate determinants of fertility

John Bongaarts \*

[Show more](#) ▾

[https://doi.org/10.1016/0160-791X\(87\)90003-0](https://doi.org/10.1016/0160-791X(87)90003-0)





[Get rights and content](#)

## Age patterns of mortality for older women: An analysis using the age-specific rate of mortality change with age

Shiro Horiuchi & Ansley J. Coale

### RESEARCH ARTICLE

## Tracking the reach of COVID-19 kin loss with a bereavement multiplier applied to the United States

 Ashton M. Verdery,  Emily Smith-Greenaway,  Rachel Margolis, and  Jonathan Daw





# Motivation

## Preferences, Partners, and Parenthood: Linking Early Fertility Desires, Marriage Timing, and Achieved Fertility

Natalie Nitsche & Sarah R. Hayford

OriginalPaper | Open Access | Published: 12 November 2020 | Pages: 1975 - 2001



U.S.

## The Dynamics of Intimate Relationships and Contraceptive Use During Early Adulthood

Yasamin Kusunoki & Jennifer S. Barber

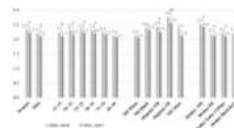
OriginalPaper | Published: 08 September 2020 | Pages: 2003 - 2034

U.S.

## Recent Trends in U.S. Childbearing Intentions

Caroline Sten Hartnett & Alison Gemmill

OriginalPaper | Published: 09 November 2020 | Pages: 2035 - 2045



## The Enduring Case for Fertility Desires

Sara Yeatman, Jenny Trinitapoli & Sarah Garver

OriginalPaper | Published: 01 October 2020 | Pages: 2047 - 2056



Malawi

## Material Hardship and Contraceptive Use During the Transition to Adulthood

Eily Field

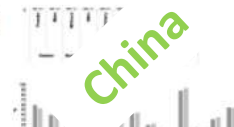
OriginalPaper | Published: 30 November 2020 | Pages: 2057 - 2084

U.S.

## Parental Investment After the Birth of a Sibling: The Effect of Family Size in Low-Fertility China

Shuang Chen

OriginalPaper | Published: 29 October 2020 | Pages: 2085 - 2111



China

## Exposure to Armed Conflict and Fertility in Sub-Saharan Africa

Brian Colman, Matthew Hancock ... James Piazza

OriginalPaper | Published: 16 October 2020 | Pages: 2113 - 2141

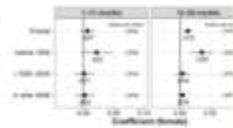


Africa

## Gender Discrimination and Excess Female Under-5 Mortality in India: A New Perspective Using Mixed-Sex Twins

Neelam K. Behman

OriginalPaper | Open Access | Published: 25 September 2020 | Pages: 2143 - 2167



## The Effects of Marital Status, Fertility, and Bereavement on Adult Mortality in Polygamous and Monogamous Households: Evidence From the Utah Population Database

Kieron J. Barclay, Robyn Donovick Thorén ... Rebecca L. Hanson

OriginalPaper | Open Access | Published: 15 September 2020 | Pages: 2169 - 2190



## Landfall After the Perfect Storm: Cohort Differences in the Relationship Between Debt and Risk of Heart Attack

Angela M. O'Rand & Jennifer Ham-Luker

OriginalPaper | Published: 13 October 2020 | Pages: 2199 - 2220



U.S.

## Young Adults' Migration to Cities in Sweden: Do Siblings Pave the Way?

Clara H. Mulder, Emma Lundholm & Gunnar Malmberg

OriginalPaper | Open Access | Published: 30 November 2020 | Pages: 2221 - 2244



## Evaluating the Role of Parental Education and Adolescent Health Problems in Educational Attainment

Janne Mäkitinen, Hanna Remes ... Pekka Martikainen

OriginalPaper | Open Access | Published: 01 October 2020 | Pages: 2245 - 2267

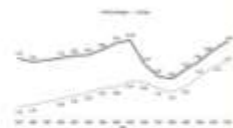


Finland

## Historical Trends in Children Living in Multigenerational Households in the United States: 1870-2018

Nafesa V. Piliakos, Mariana Almonst & Rachel E. Dunifon

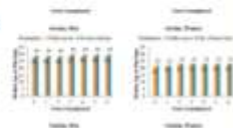
OriginalPaper | Published: 01 October 2020 | Pages: 2269 - 2296



## Employment's Role in Enabling and Constraining Marriage in the Middle East and North Africa

Caroline Kraft & Ragui Assaad

OriginalPaper | Published: 29 October 2020 | Pages: 2297 - 2323



## The Association Between Legal Status and Poverty Among Immigrants: A Methodological Caution

Cody Spence, James D. Bachmeier ... Christal Hamilton

OriginalPaper | Published: 29 October 2020 | Pages: 2327 - 2335

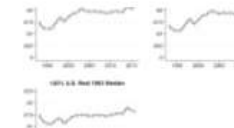


U.S.

## The Levels and Trends in Deep and Extreme Poverty in the United States, 1993-2016

David Brady & Zachary Parolin

OriginalPaper | Published: 15 October 2020 | Pages: 2337 - 2360



## A Cautionary Tale of Using Data From the Tail

James X. Sullivan

Commentary | Published: 15 October 2020 | Pages: 2361 - 2368



U.S.

## Further Analyses Reinforce Our Conclusions About Extreme Poverty

David Brady & Zachary Parolin

Commentary | Published: 15 October 2020 | Pages: 2369 - 2376



U.S.

## Another Plea for Caution When Using Survey Income Data From the Far

James X. Sullivan

Commentary | Published: 19 November 2020 | Pages: 2377 - 2381



U.S.

## Acknowledgment of Reviewers

Acknowledgments | Published: 30 September 2020 | Pages: 2383 - 2388

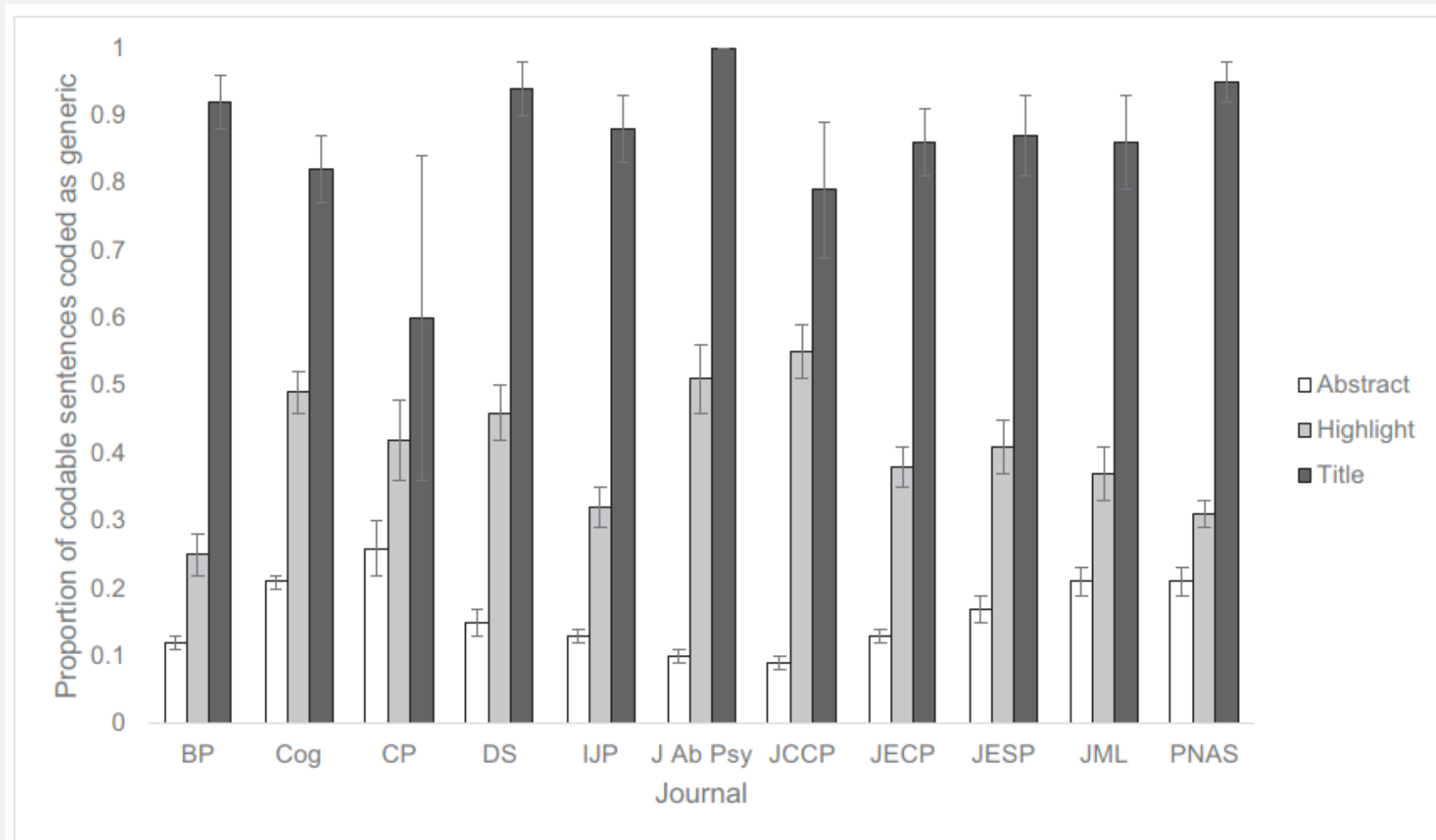


# Why focus on country names in titles?

- Most visible part of article: summarize and attract readers
- Titles with no geographical references offer less information
- Delocalized titles are more generic than localized titles
- Generic statements can be misleading (e.g., “boys are different than girls”)
- Presumption of universality!

# Generic language more common in titles

Proportion of sentences in titles, highlights, and abstracts coded as generic (number of generic sentences divided by the number of codable sentences for that component, to derive a percentage), separately by journal.



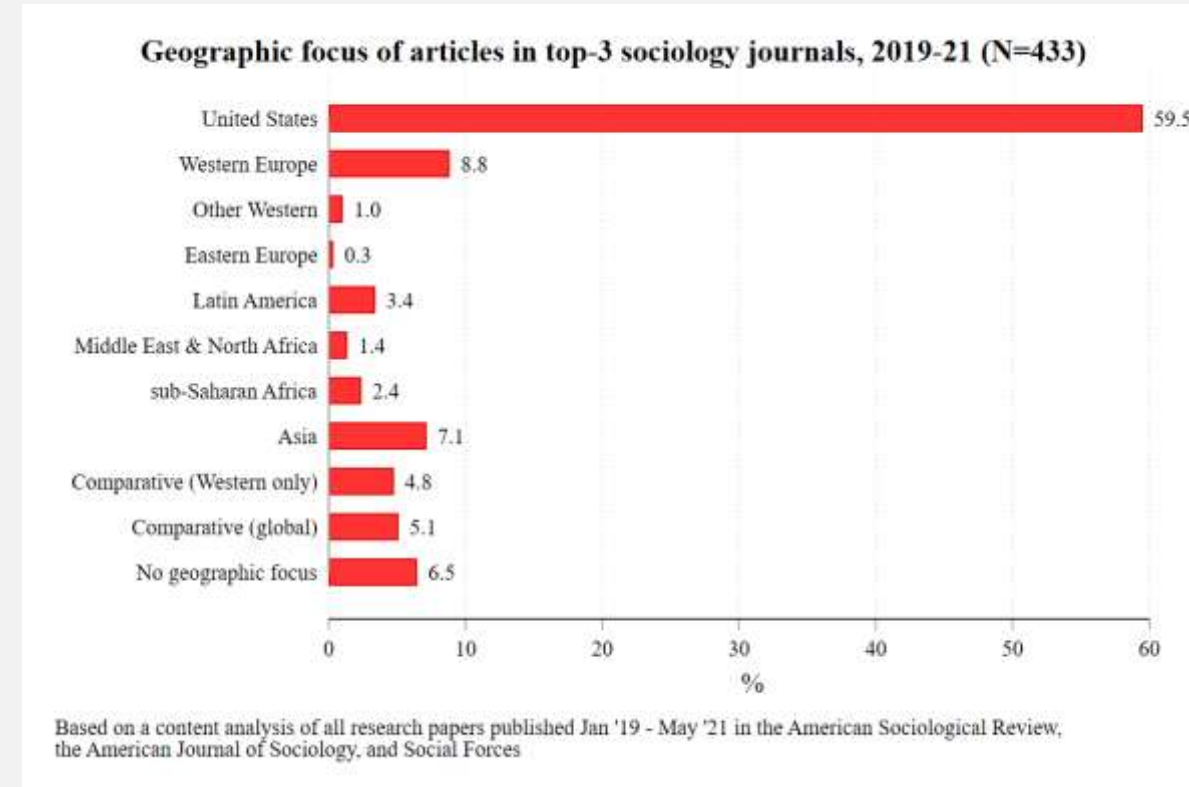
Source: DeJesus, J. M., Callanan, M. A., Solis, G. & Gelman, S. A. Generic language in scientific communication. *Proc Natl Acad Sci USA* **116**, 18370–18377 (2019).



# Eurocentrism in Social Sciences

“Social Science has been Eurocentric throughout its institutional history, [...]. Social science is a product of the modern world-system, and Eurocentrism is constitutive of the geoculture of the modern world”

Wallerstein I (1997) Eurocentrism and its Avatars: The Dilemmas of Social Science. *Sociological Bulletin* 46(1): 21–39. DOI: [10.1177/0038022919970102](https://doi.org/10.1177/0038022919970102).



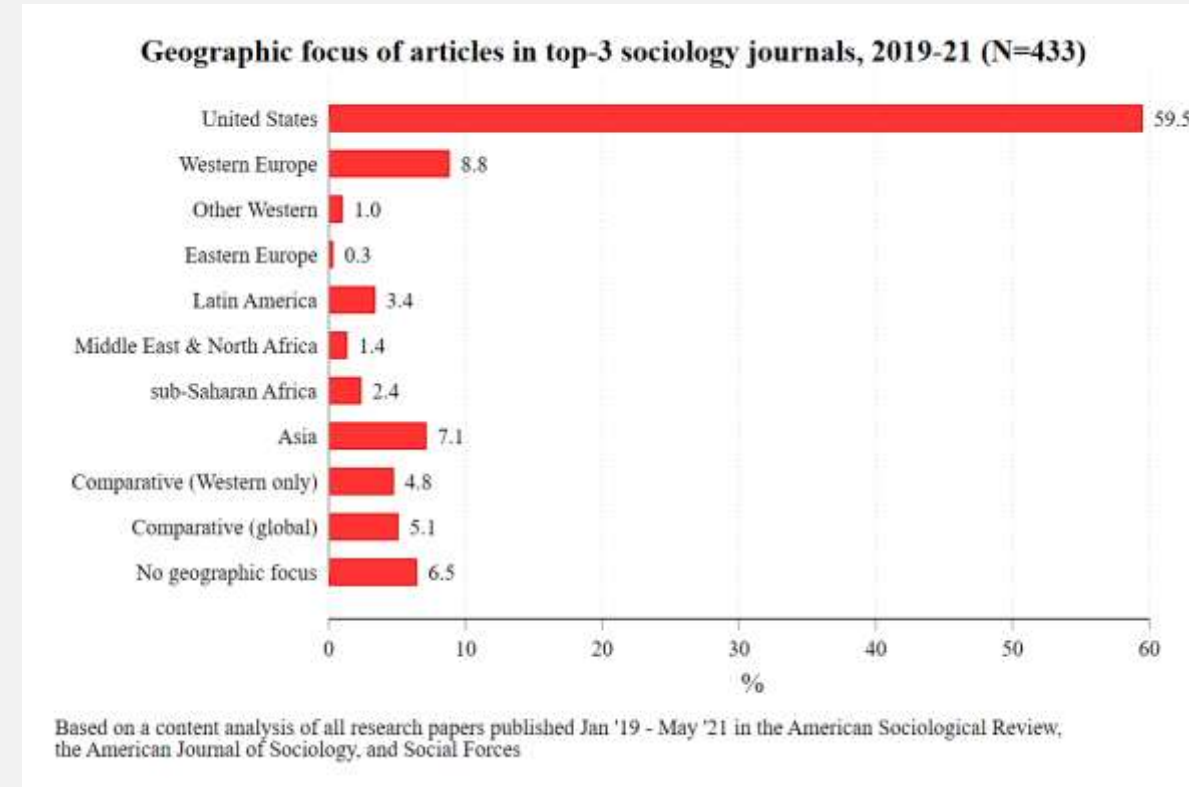
Taken from Rob Gruijters' Twitter account



# The West as the reference case

“The unsaid privileging of Western countries and Western institutions as **reference cases** is reinforced via Anglo-American journals, which allow some cases in as the default and ask those writing on other cases to explain themselves in much more details.”


Krause M (2016) ‘Western hegemony’ in the social sciences: fields and model systems: ‘Western hegemony’ in the social sciences: fields and model systems. *The Sociological Review Monographs* 64(2): 194–211. DOI: [10.1002/2059-7932.12008](https://doi.org/10.1002/2059-7932.12008).



Taken from Rob Gruijters' Twitter account

RESEARCH ARTICLE | SOCIAL SCIENCES |    

# North and South: Naming practices and the hidden dimension of global disparities in knowledge production

Andrés F. Castro Torres   and Diego Alburez-Gutierrez  [Authors Info & Affiliations](#)

Edited by Susan Fiske, Princeton University, Princeton, NJ; received October 27, 2021; accepted January 23, 2022

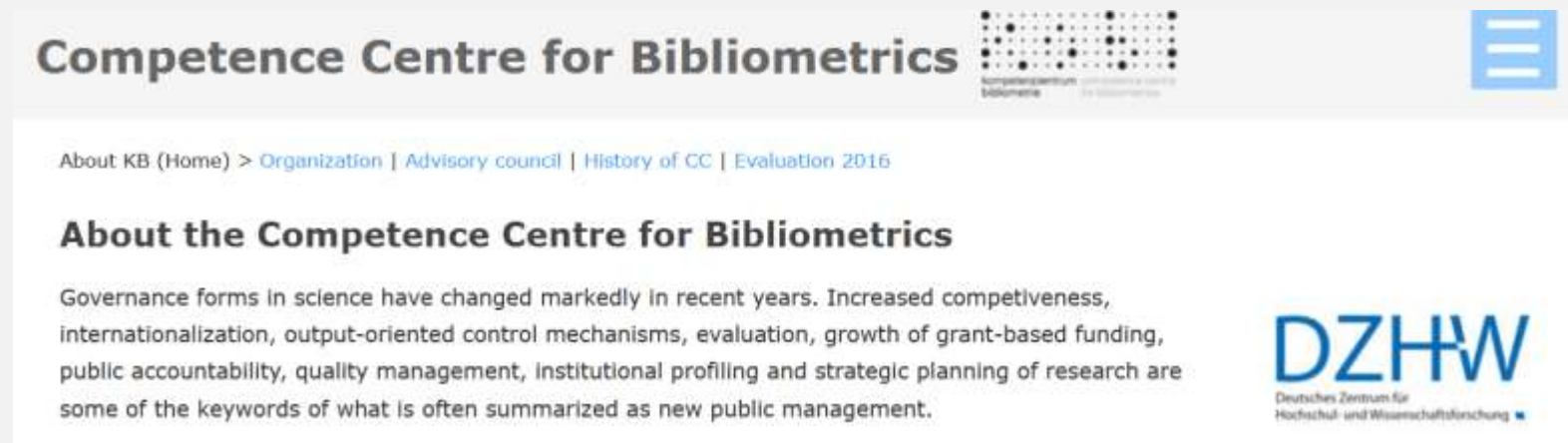
**March 1, 2022** | 119 (10) e2119373119 | <https://doi.org/10.1073/pnas.2119373119>

# Data and Methods



# Data

- Scopus-indexed publication records (ca. 80 million)
- SQL relational database
- Title, abstract, classification, authors, year, journal, affiliation, etc.
- Competence Centre for Bibliometrics



The screenshot shows the homepage of the Competence Centre for Bibliometrics. The header features the organization's name, a logo of a grid of dots, and a blue menu icon. A navigation bar contains links: 'About KB (Home)', 'Organization', 'Advisory council', 'History of CC', and 'Evaluation 2016'. The main content area is titled 'About the Competence Centre for Bibliometrics' and contains a paragraph about governance in science. The DZHW logo is in the bottom right corner.

**Competence Centre for Bibliometrics**

[About KB \(Home\)](#) > [Organization](#) | [Advisory council](#) | [History of CC](#) | [Evaluation 2016](#)

### About the Competence Centre for Bibliometrics

Governance forms in science have changed markedly in recent years. Increased competitiveness, internationalization, output-oriented control mechanisms, evaluation, growth of grant-based funding, public accountability, quality management, institutional profiling and strategic planning of research are some of the keywords of what is often summarized as new public management.

**DZHW**  
Deutsches Zentrum für  
Hochschul- und Wissenschaftsforschung

# Sample selection

- Empirical studies only (remove “chapter,” “editorial,” “book review,” “erratum”, etc.)
- Published between 1996 to 2020
- In “Social Sciences and Humanities” (27 subfields of ACJS category)
- Have an abstract
- Mention at least one country in the abstract

**Sample size:**





# Social science subfields

Code ▾	Field
1212	Religious Studies
1410	Industrial Relations
2000	General Economics, Econometrics and Finance
2001	Economics, Econometrics and Finance (miscellaneous)
2002	Economics and Econometrics
3200	General Psychology
3201	Psychology (miscellaneous)
3202	Applied Psychology
3203	Clinical Psychology
3204	Developmental and Educational Psychology
3205	Experimental and Cognitive Psychology
3207	Social Psychology
3300	General Social Sciences
3301	Social Sciences (miscellaneous)
3303	Development
3304	Education
3305	Geography, Planning and Development
3306	Health (social science)
3312	Sociology and Political Science
3314	Anthropology
3316	Cultural Studies
3317	Demography
3318	Gender Studies
3319	Life-span and Life-course Studies
3320	Political Science and International Relations
3321	Public Administration
3322	Urban Studies

Code ▾	Field
1200	General Arts and Humanities
1201	Arts and Humanities (miscellaneous)
1202	History
1203	Language and Linguistics
1204	Archeology (arts and humanities)
1205	Classics
1206	Conservation
1207	History and Philosophy of Science
1208	Literature and Literary Theory
1209	Museology
1210	Music
1211	Philosophy
1213	Visual Arts and Performing Arts
1400	General Business, Management and Accounting
1401	Business, Management and Accounting (miscellaneous)
1402	Accounting
1403	Business and International Management
1404	Management Information Systems
1405	Management of Technology and Innovation
1406	Marketing
1407	Organizational Behavior and Human Resource Management
1408	Strategy and Management
1409	Tourism, Leisure and Hospitality Management
1800	General Decision Sciences
1801	Decision Sciences (miscellaneous)
1802	Information Systems and Management
1803	Management Science and Operations Research
1804	Statistics, Probability and Uncertainty
2003	Finance
3206	Neuropsychology and Physiological Psychology
3302	Archeology
3307	Human Factors and Ergonomics
3308	Law
3309	Library and Information Sciences
3310	Linguistics and Language
3311	Safety Research
3313	Transportation

# Measure of interest

$$\textit{Localization Rate (LR)} = \frac{n(\textit{title includes a country} \mid \textit{abstract includes a country})}{n(\textit{abstract includes a country})}$$



# Sample

We classify articles according to the inclusion (or lack thereof) of a country name or country demonym (e.g., Colombian) in the title and abstract. We focus on papers with at least one country mentioned in the abstract.

At least one country mentioned in:			
Article's abstract	Article's title		Total
	No	Yes	
No	656,235	30,652	686,887
Yes	324,217	331,520	655,737
No abstract	201,949	39,386	241,335
Total	1,182,401	401,558	1,583,959

*Note:* The data include articles from 27 fields of the “Social Sciences & Humanities” category, indexed by the Web of Science and SCOPUS, only publications in English are included.



# Sample

We classify articles according to the inclusion (or lack thereof) of a country name or country demonym (e.g., Colombian) in the title and abstract. We focus on papers with at least one country mentioned in the abstract.

At least one country mentioned in:			
Article's abstract	Article's title		Total
	No	Yes	
No	656,235	30,652	686,887
Yes	324,217	331,520	655,737
No abstract	201,949	39,386	241,335
Total	1,182,401	401,558	1,583,959

Localized papers

*Note:* The data include articles from 27 fields of the “Social Sciences & Humanities” category, indexed by the Web of Science and SCOPUS, only publications in English are included.



# Sample

We classify articles according to the inclusion (or lack thereof) of a country name or country demonym (e.g., Colombian) in the title and abstract. We focus on papers with at least one country mentioned in the abstract.

At least one country mentioned in:			
Article's abstract	Article's title		Total
	No	Yes	
No	656,235	30,652	686,887
Yes	324,217	331,520	655,737
	49%	51%	100%
No abstract	201,949	39,386	241,335
Total	1,182,401	401,558	1,583,959

Localization Rate (LR)

*Note:* The data include articles from 27 fields of the “Social Sciences & Humanities” category, indexed by the Web of Science and SCOPUS, only publications in English are included.



# Sample

We classify articles according to the inclusion (or lack thereof) of a country name or country demonym (e.g., Colombian) in the title and abstract. We focus on papers with at least one country mentioned in the abstract.

At least one country mentioned in:			
Article's abstract	Article's title		Total
	No	Yes	
No	656,235	30,652	686,887
	96%	4%	100%
Yes	324,217	331,520	655,737
	49%	51%	100%
No abstract	201,949	39,386	241,335
	84%	16%	100%
Total	1,182,401	401,558	1,583,959
	75%	25%	100%

*Note:* The data include articles from 27 fields of the “Social Sciences & Humanities” category, indexed by the Web of Science and SCOPUS, only publications in English are included.



# Country name extraction

- Assume that country name reflects geographical focus of paper
- Locate country name in title and abstract
- Exclude subnational or supranational entities (e.g., cities, continents)
- Exclude copyright statements (e.g., “© Akadémiai Kiado, Budapest, Hungary 2014”).



# Some examples from our data set

## Localized papers

- [1] "Demographic transition, demographic dividend and economic growth in Nigeria"
- [2] "The meeting of myths and realities: The \"homecoming\" of second-generation exiles in post-apartheid South Africa"
- [3] "Untraditional caring arrangements among parents living apart: The case of Norway"
- [4] "Human Rights, Positive Obligations, and Measures to Prevent Human Trafficking in the United Kingdom"
- [5] "A new projection of the number of recipients of the care allowance to the year 2030 in the Czech Republic"
- [6] "Fertility and education among British Asian women: A success story of social mobility?"

## Non-localized papers

- [1] "International labor visits and knowledge transfers"
  - [2] "Cohort change, diffusion, and support for gender egalitarianism in cross-national perspective"
  - [3] "Adoption quarterly special double issue: Using data from the national survey of adoptive parents"
  - [4] "HOW DOES THE WELFARE STATE INFLUENCE INDIVIDUALS' SOCIAL CAPITAL?: Eurobarometer evidence on individuals' access to informal help"
  - [5] "The making and unmaking of precarious, ideal subjectsâ€"migration brokerage in the Global South"
  - [6] "Intermediaries and destination reputations: explaining flows of skilled migration"
- > |

# Regular expression

- ISO 3166–1 country names, abbreviations and demonyms (“countrycode” R package)

```
\bnigeria\b|\bnigerian\b
```

```
\bhungary\b|\bhungarian\b|\bmagyar\b
```

```
^(?!.*\bdem) (?!.*\bd[\.]?r) (?!.*kinshasa) (?!.*za  
ire) (?!.*belg) (?!.*l.opoldville) (?!.*free) .* \bco  
ngo
```

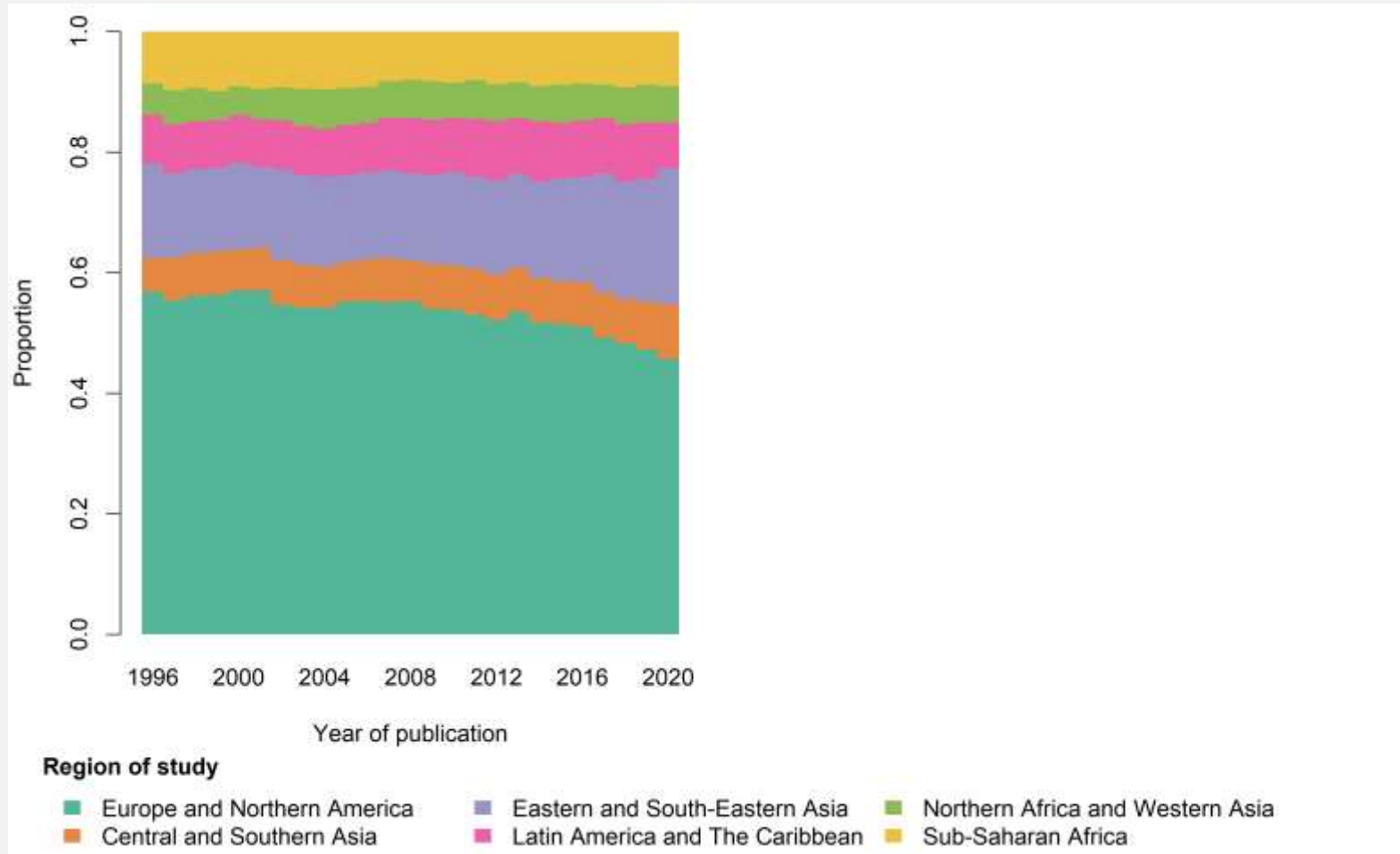
# Results



# Science production by region



**Fig. 2.** Distribution of the region of focus of articles in the analytical sample (left) and localization rate by region of study (right), 1996-2020 (n = 560,893).



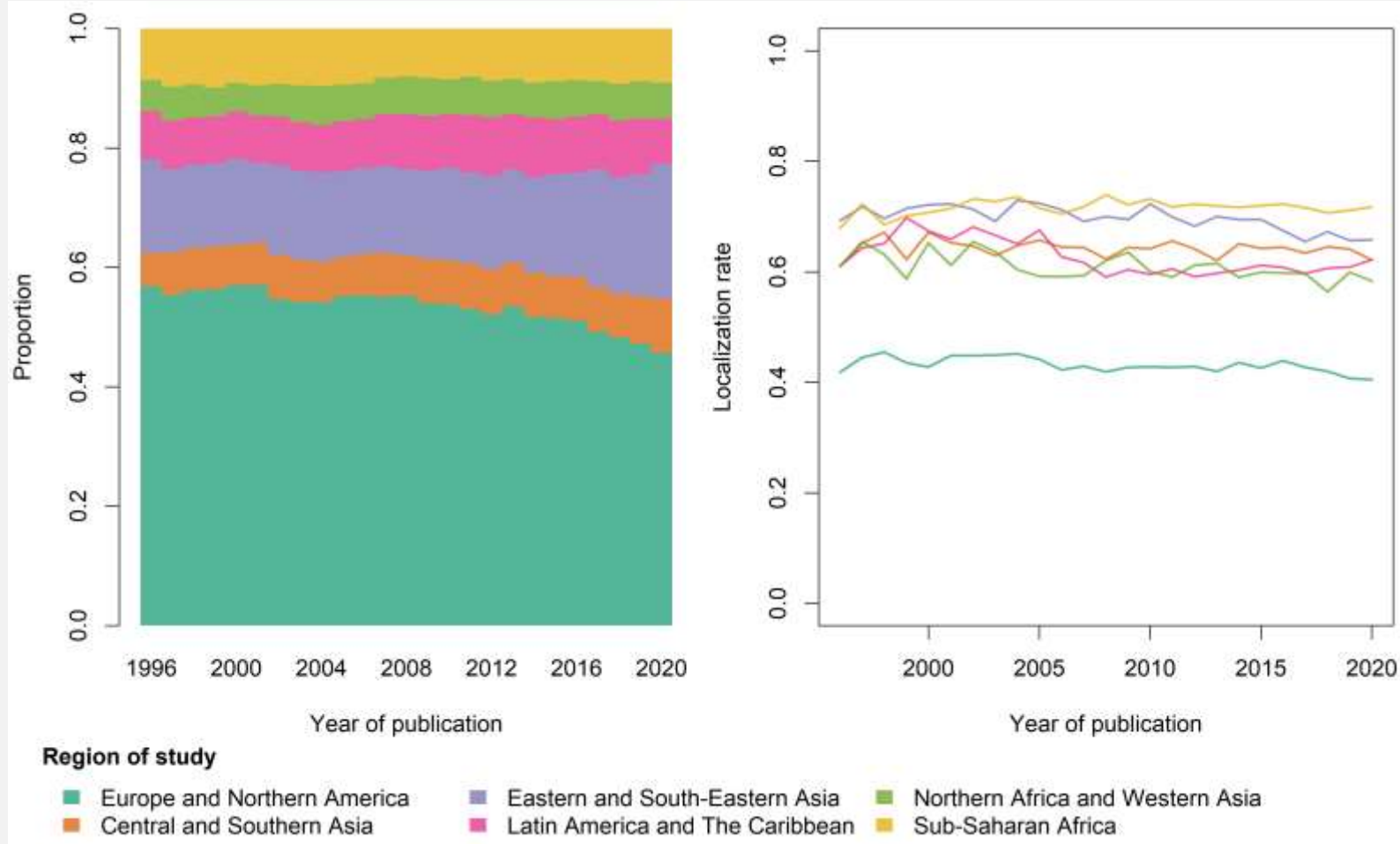
Note: The regional classification of countries is taken from the United Nations. The category “Europe and North America” includes Australia and New Zealand. Papers about Oceania are excluded from the analysis due to the small sample size (n = 1,583).



# Science production AND localization rates



**Fig. 2.** Distribution of the region of focus of articles in the analytical sample (left) and localization rate by region of study (right), 1996-2020 (n = 560,893).

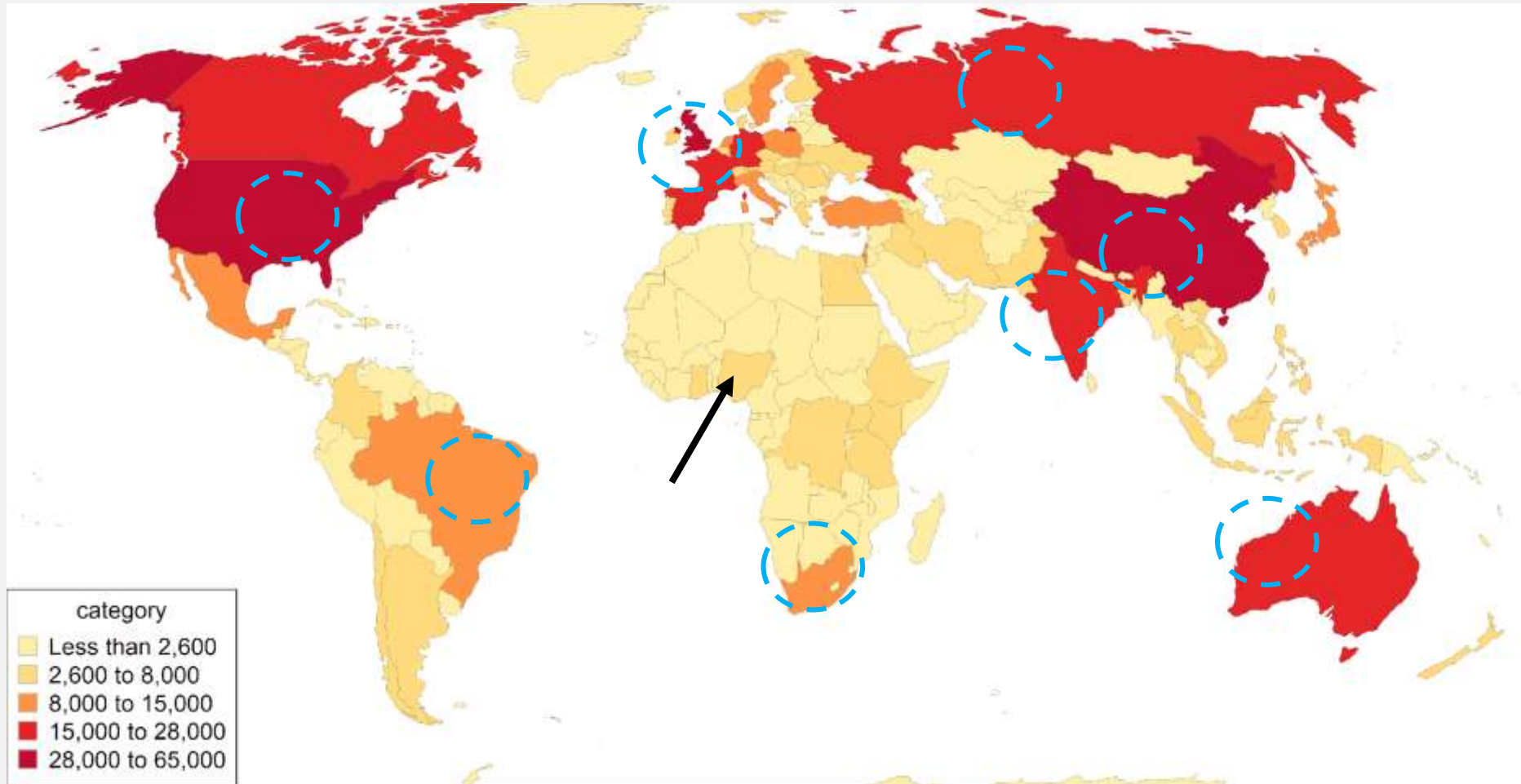


Note: The regional classification of countries is taken from the United Nations. The category “Europe and North America” includes Australia and New Zealand. Papers about Oceania are excluded from the analysis due to the small sample size (n = 1,583).



# Distribution of science production

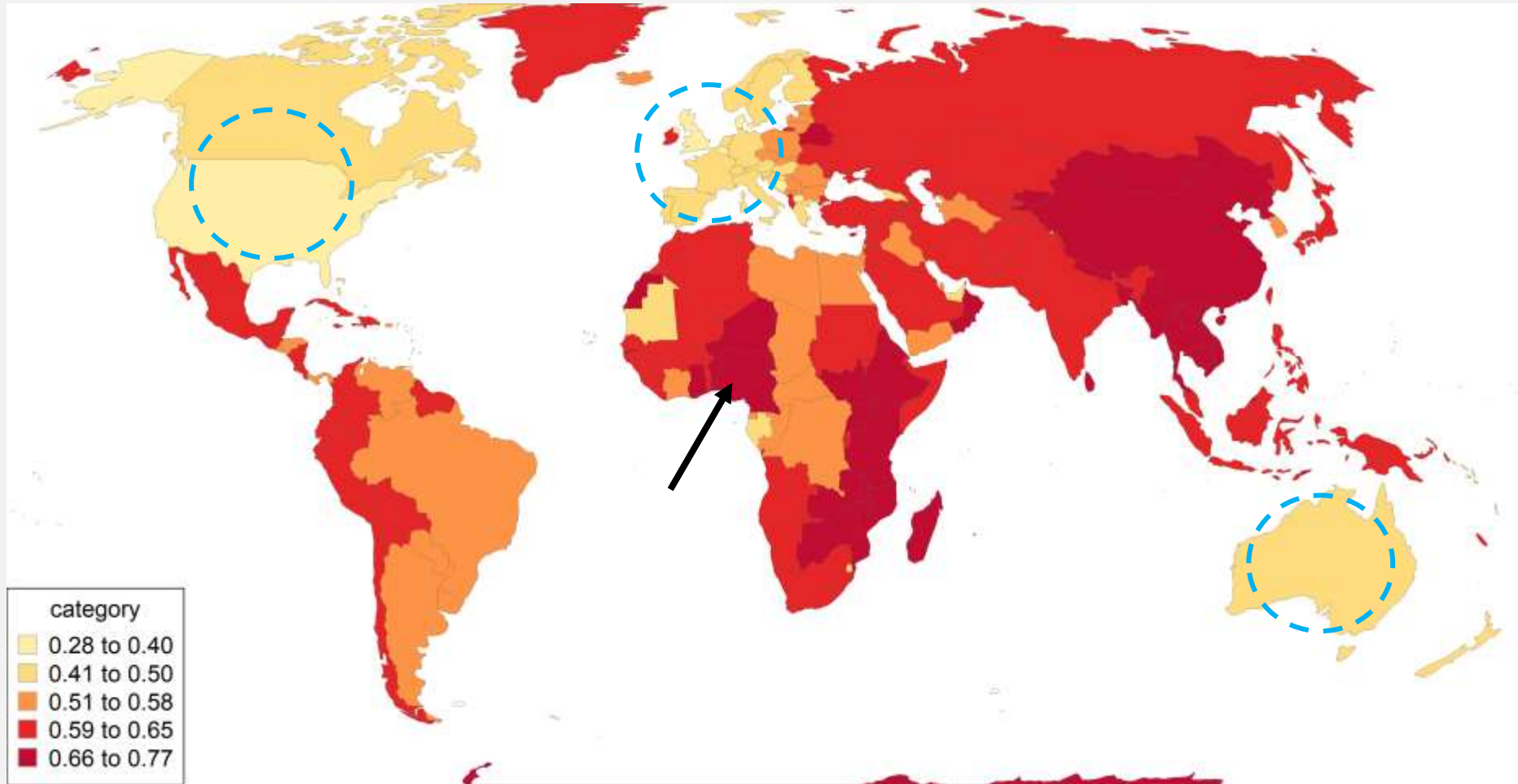
Number of mentions in articles' abstract in Social Sciences and Humanities, 1996-2020



*Note:* The data include articles from 27 fields of the “Social Sciences & Humanities” category, indexed by SCOPUS, only publications in English are included.



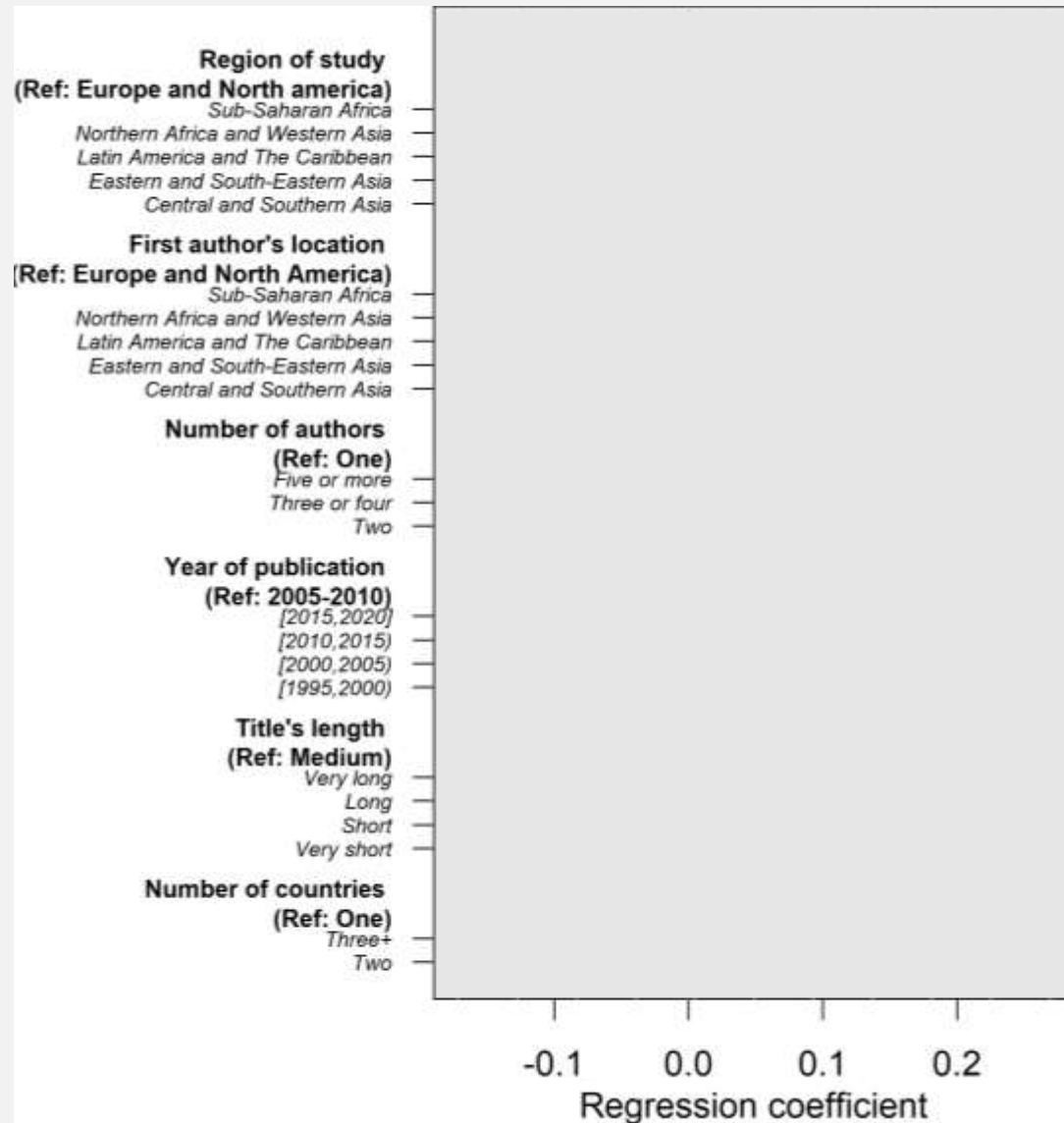
# Country-level localization rates



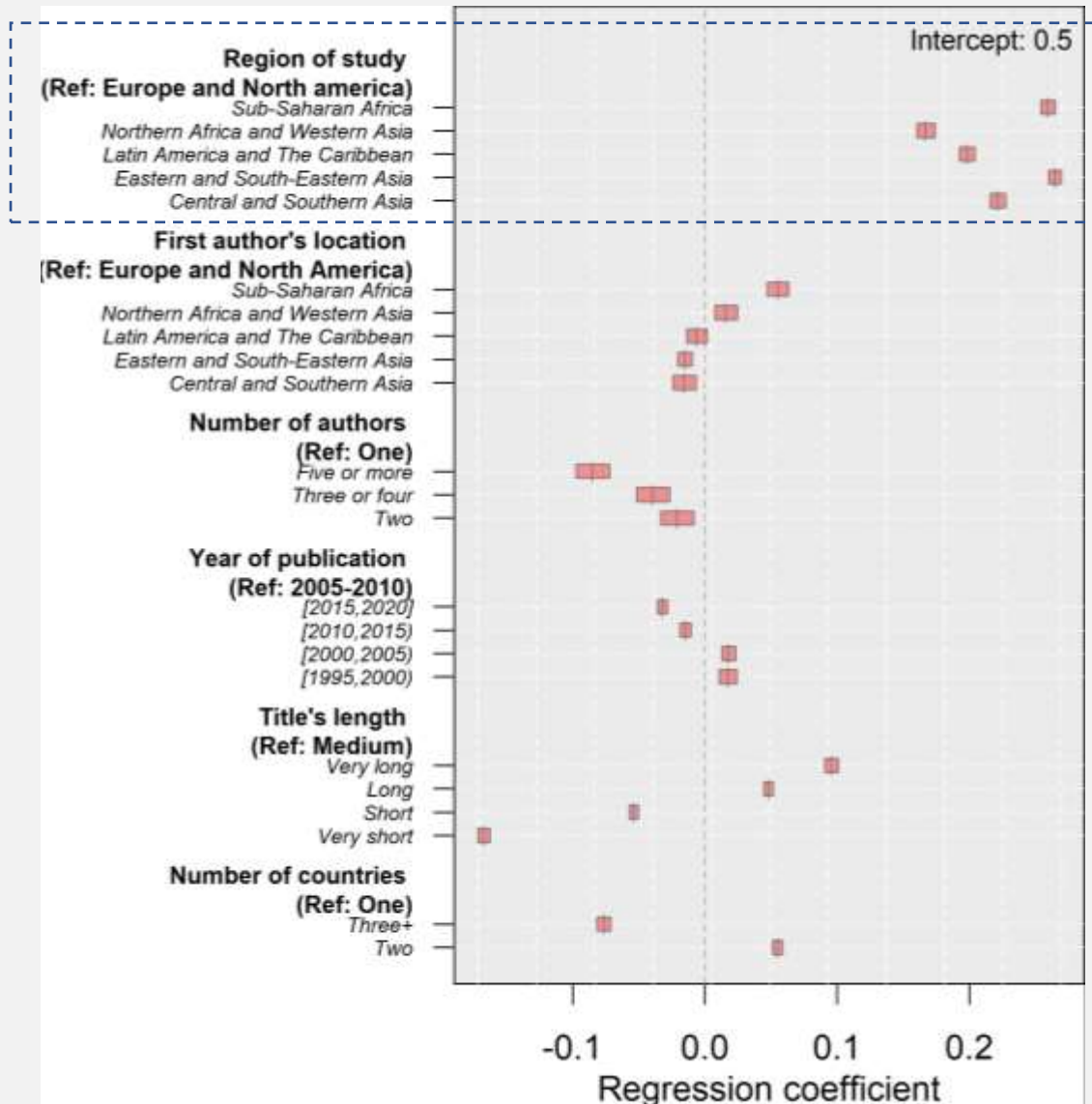
*Note:* The data include articles from 27 fields of the “Social Sciences & Humanities” category, indexed by the Web of Science and SCOPUS, only publications in English are included.



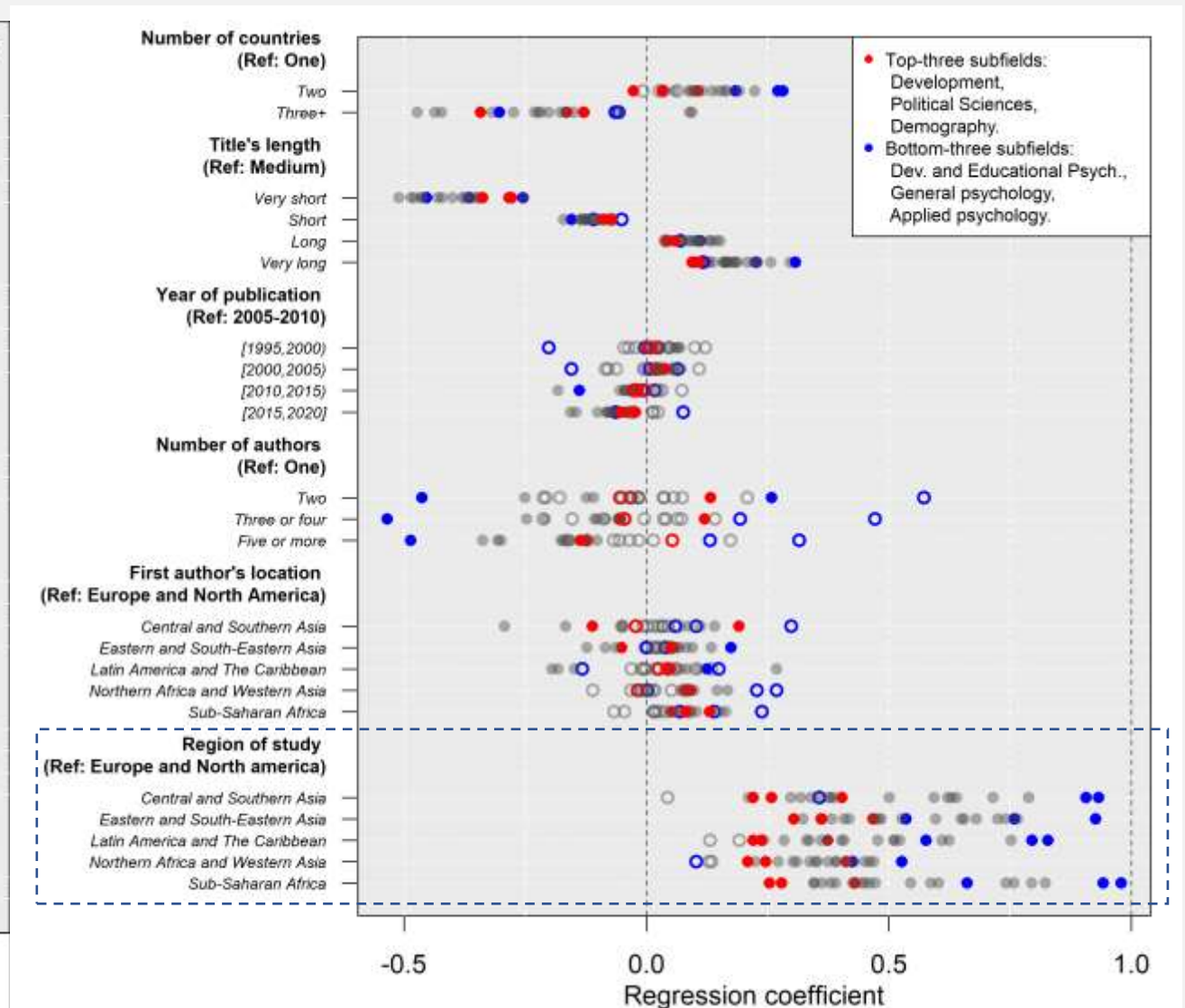
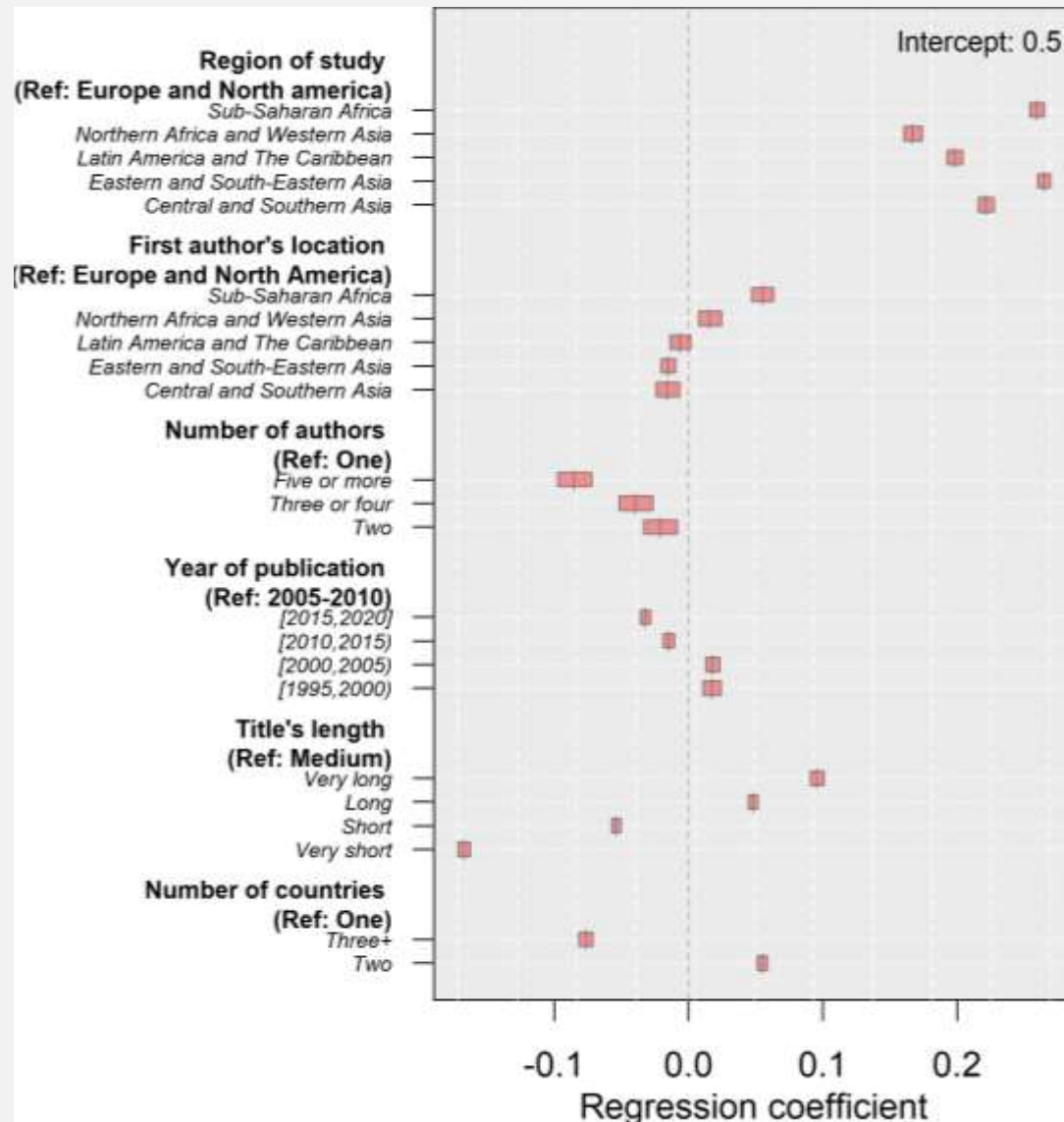
# Regression analysis



# Countries outside Global North more likely to localize

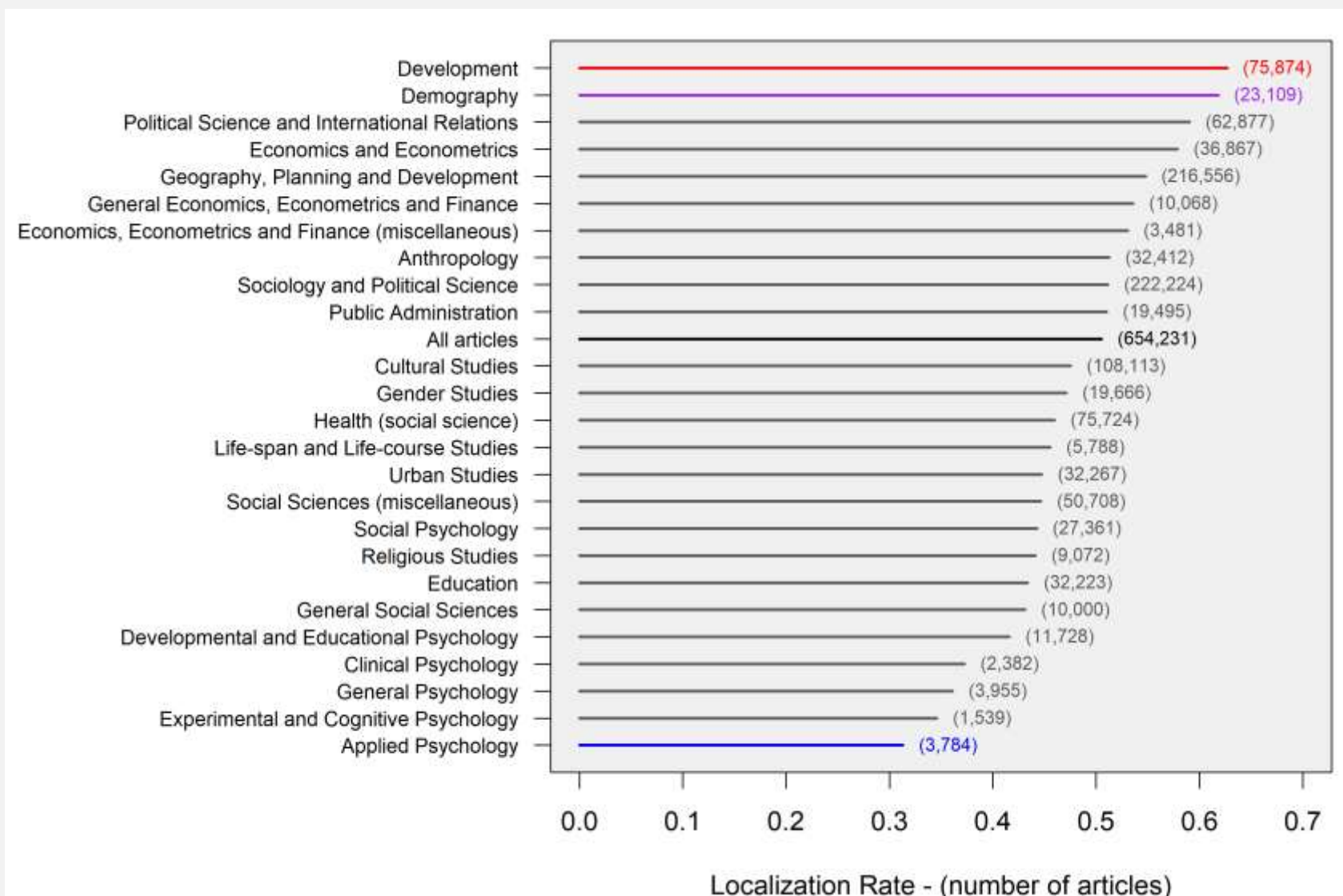


# Results hold for all subdisciplines





# Localization rate by subdiscipline



# Additional analyses

# Robustness I

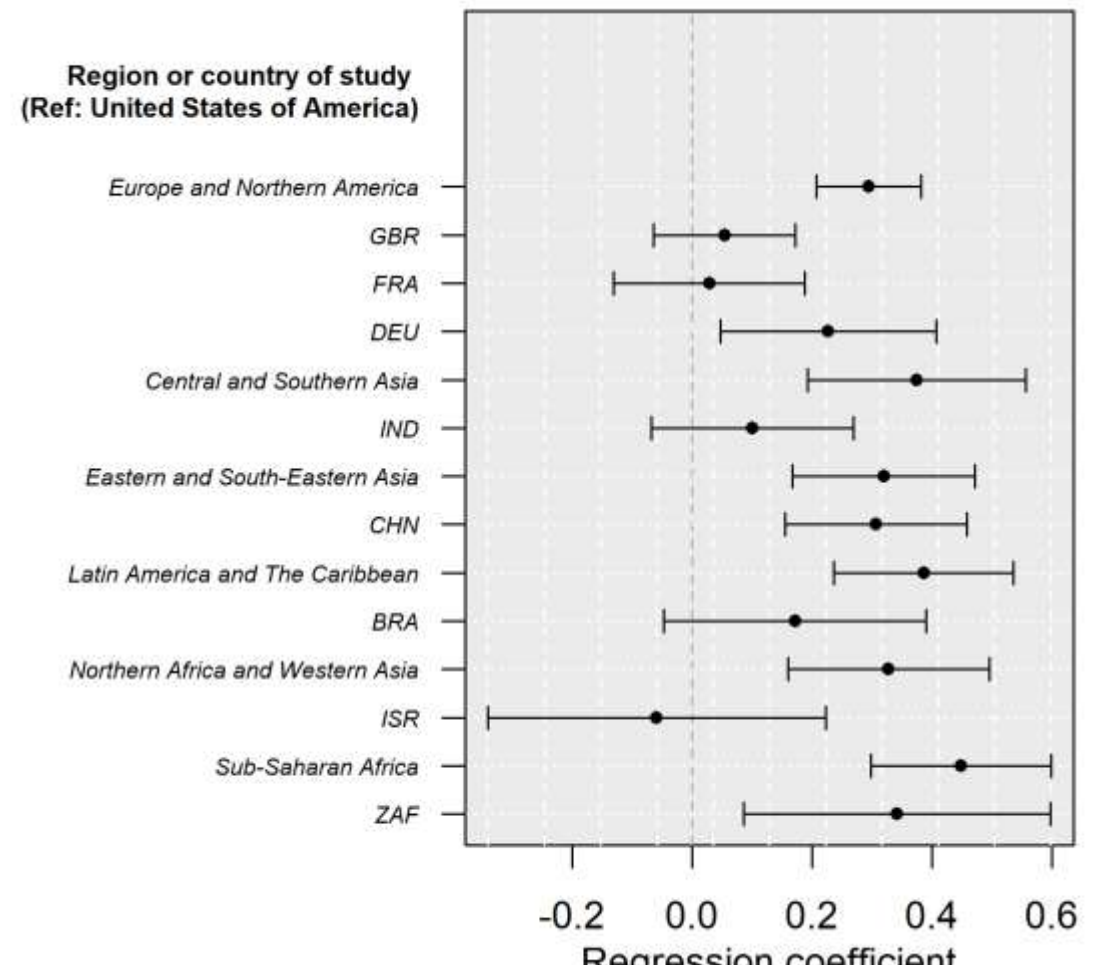
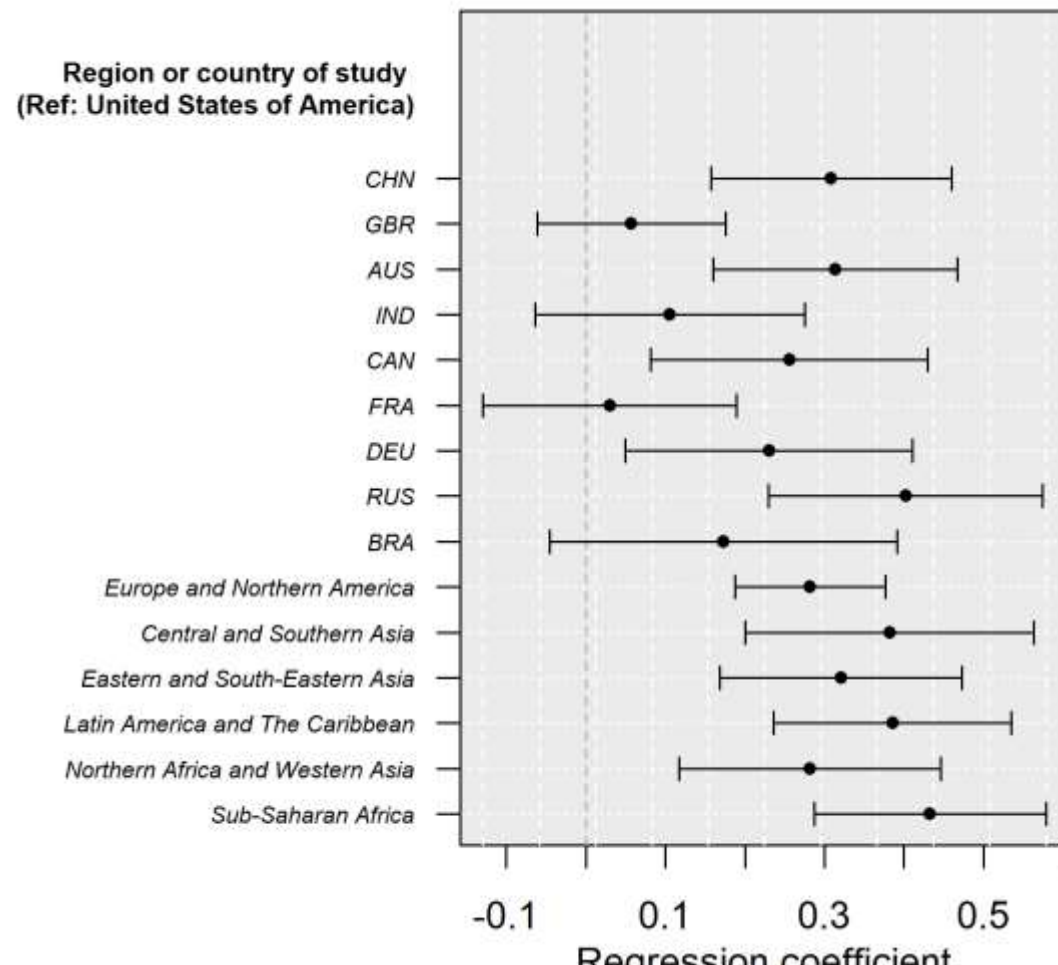
- Two RAs hand coded stratified representative random sample of Scopus data (n = 2,510).
- Research items' titles includes a reference to
  - a national entity (i.e., a country name or demonym)?
  - a subnational entity (e.g., a region or city within a country),
  - or a supranational entity (e.g., a continent)?

# Robustness I

- Algorithms successfully identified 94.3% (SE =0.39) of localized articles.
- Algorithms missed 5.2% (SE =0.42) of subnational and 4.5% (SE =0.39) of supranational units
- 9.3% (s.e. = 0.53) of articles misclassified as being social science publications.

# Robustness II

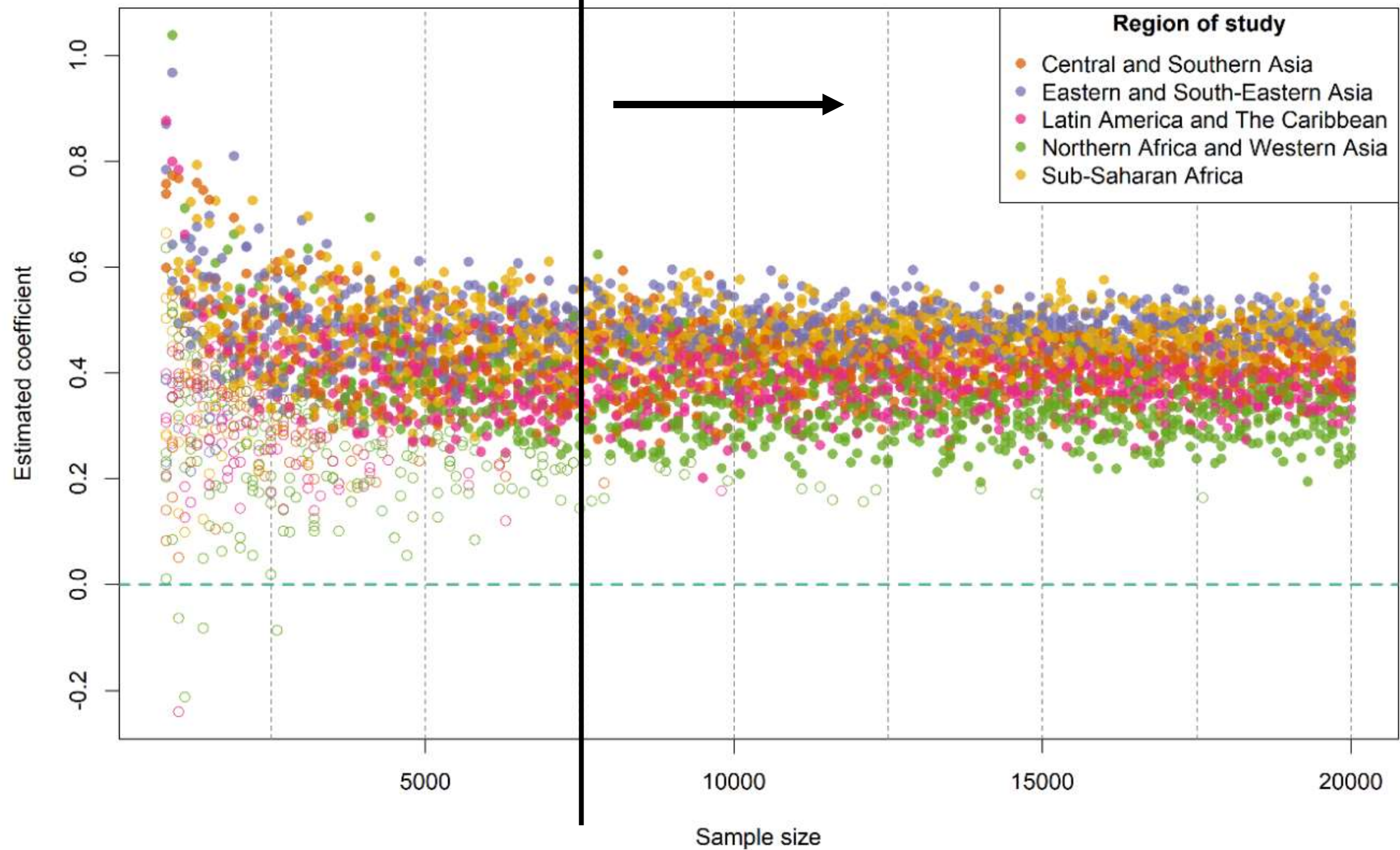
- Replicated results using hand-coded sample (1,195 papers)





# Replicating results with smaller samples

- Extracted 1,000 random samples between 800-20,000 papers
- 2,500-5,000 articles: all regions except for Northern Africa, Western Asia, Latin America, and the Caribbean.
- 5,000-7,500 articles: all except Northern Africa and Western Asia.
- 7,500 articles: replicate results for all regions



# Conclusions



# Conclusions

1. Large and consistent gaps in localization rates between global North and South
2. Region-specific hegemonies (with lower localization rates than neighbors)
3. Robust findings across 23 subfields



# Conclusions

1. Knowledge production affected by North/South power imbalances
2. Evidence produced in (and refers to) global North countries more generalizable
3. Essentialist understanding of social processes?



# Implications

“[...] generic language in scientific articles may lead readers to reach exaggerated conclusions. [...] generic sentences implied that a property was broadly true and conceptually central, and that the category expressed was stable with inherent properties”

DeJesus JM, Callanan MA, Solis G, et al. (2019) Generic language in scientific communication. *Proceedings of the National Academy of Sciences* 116(37): 18370–18377. DOI: 10.1073/pnas.1817706116.



# Implications

- Essentialist understanding of social processes
- Differentials in the readability and citations
- Transplantation/imposition of Western problems
- Non-Western societies, individuals and institutions as exotic

# Thank you!

Diego Alburez-Gutierrez

[alburezgutierrez@demogr.mpg.de](mailto:alburezgutierrez@demogr.mpg.de)

@d\_alburez





# Number of countries in the abstract

