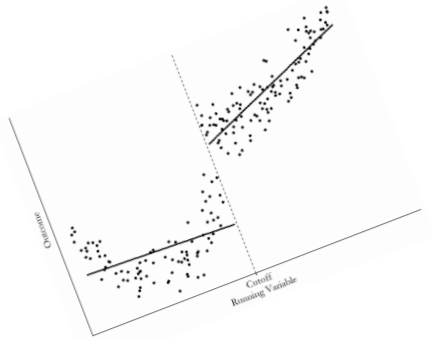
Causal Claims in Economics

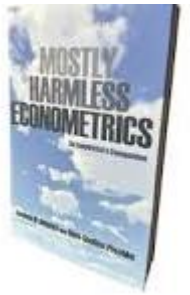
Prashant Garg
(Imperial, IFC)

Thiemo Fetzer
(Warwick, Bonn, CEPR)

www.causal.claims/

Background and Motivation





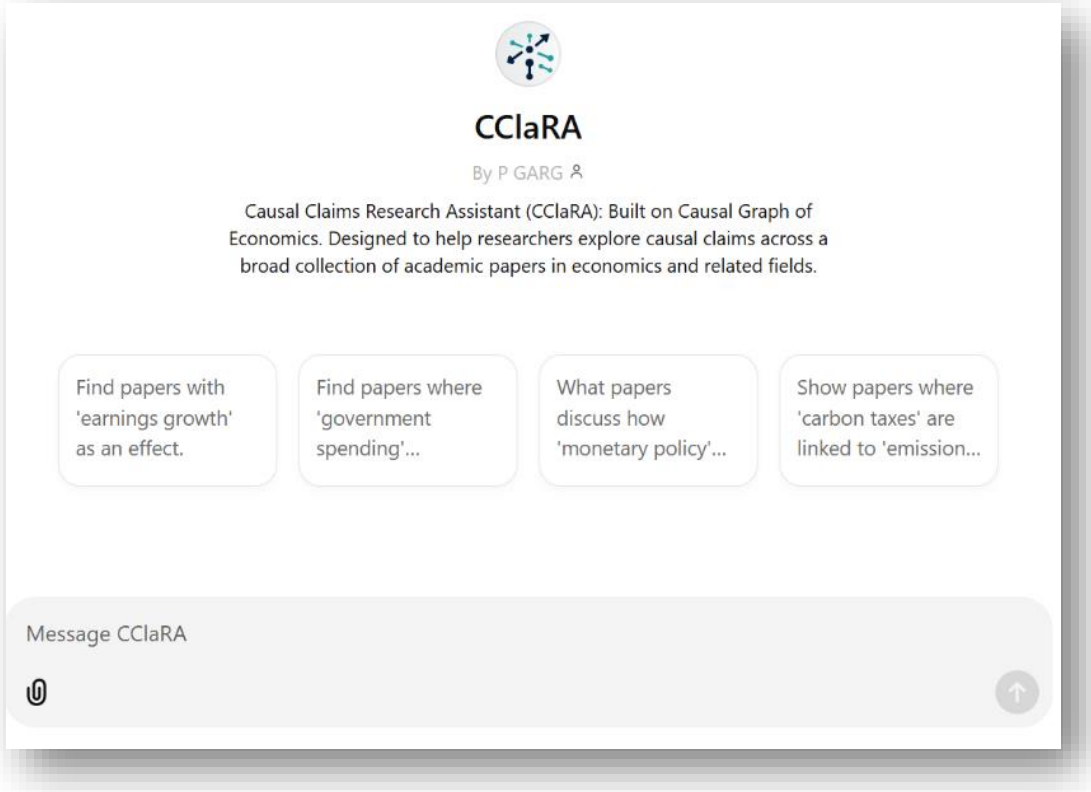
- Economics is a dynamic field that has witnessed a profound transformation over the past four decades.
- The discipline has shifted towards establishing causal relationships using advanced empirical methods—a movement known as the "credibility revolution."
- At the heart of our project is the creation of the **Causal Graph of Economics.**
- We have analyzed over 44,000 working papers from the National Bureau of Economic Research (NBER) and the Centre for Economic Policy Research (CEPR) using AI to map out the intricate network of causal claims that shape economic research.

Research Agenda

- Synthesize Causal Evidence from Economics (a global meta study) in Graphical format.
  - Website where authors can see their own causal graph: here.

- Make complex economic research accessible
  - eg, CClaRA- Causal Claims Research Assistant,

- **This paper**:
  - Description of the causal claims in Economics.



Research Questions (this paper)

1. Can we synthesise the (causal) knowledge graph in economics?
2. Description of Causal Claims
   a) How have **causal inference methods** changed over time and across different fields?
   b) Each paper has an implicit causal graph, a narrative. Has **narratives complexity** changed over time and across fields?
3. **Evaluation**:
   - How do the structure and complexity of research narratives influence publication and citation outcomes?
   - What challenges exist in replication and data accessibility?

Corpus of elite economics research

- Build a corpus of 44,000 acadmic working papers circulated as CEPR or NBER working papers
- Covering time period 1980-2024
  - 28K from **NBER**
  - 16K from **CEPR**
- Metadata from NBER, CEPR

- Citations and Publication: **RePec** + **OpenAlex**.

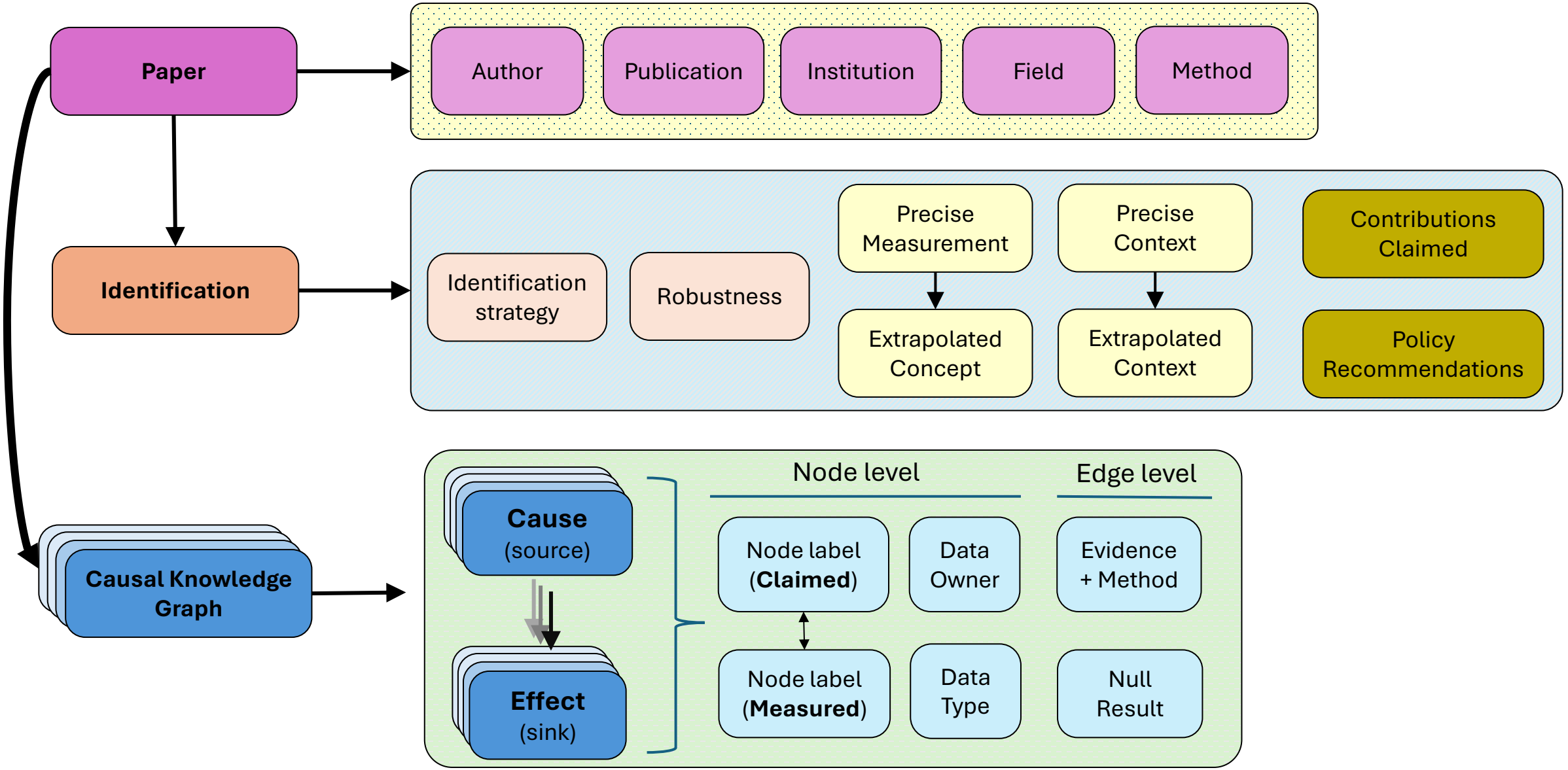
NBER
National Bureau of Economic Research

CEPR

RePEc

OpenAlex

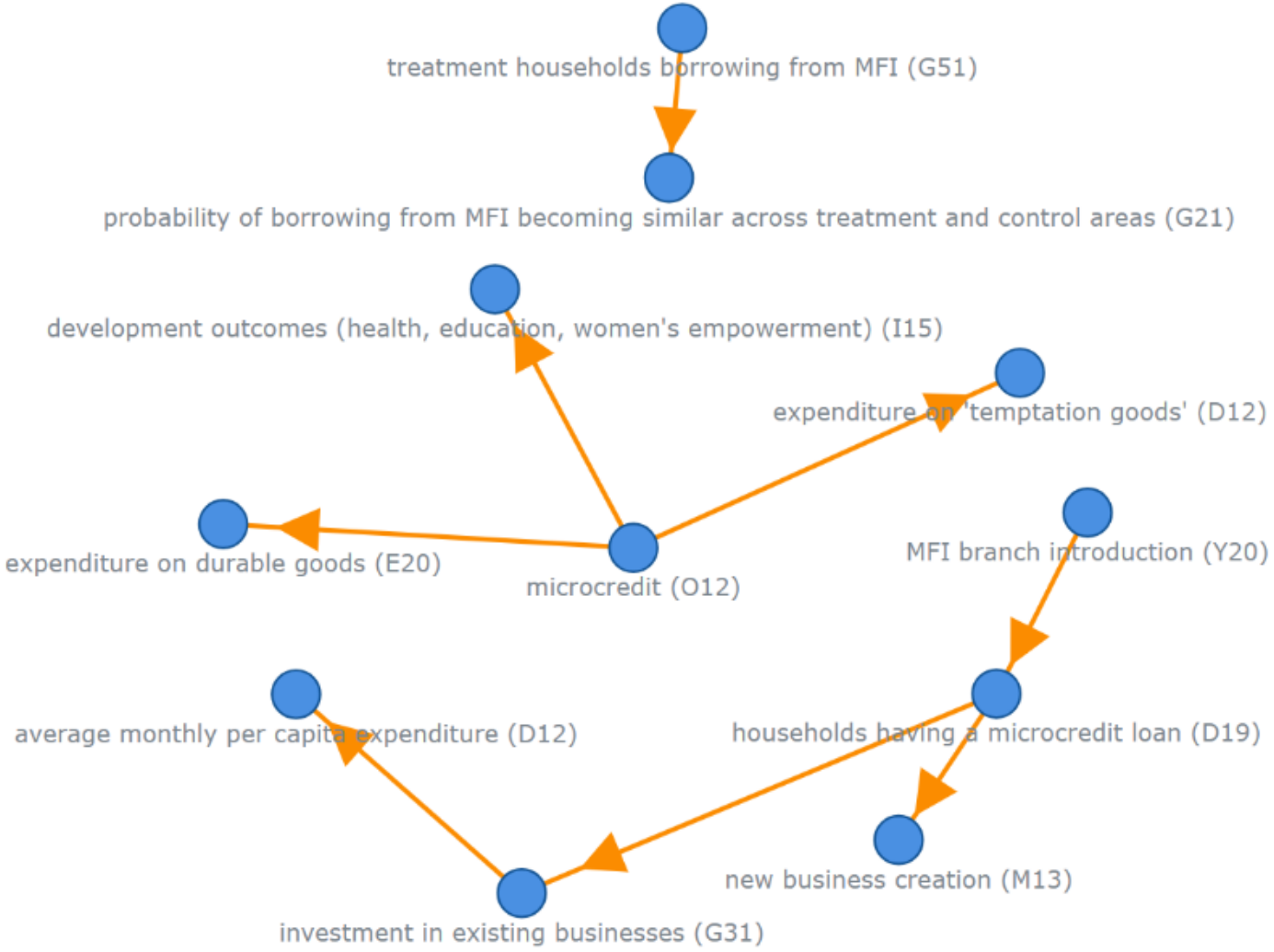
Retrieval of range of concepts using AI



An example

The Miracle of Microfinance: Evidence from a Randomized Evaluation (2015, AEJ)

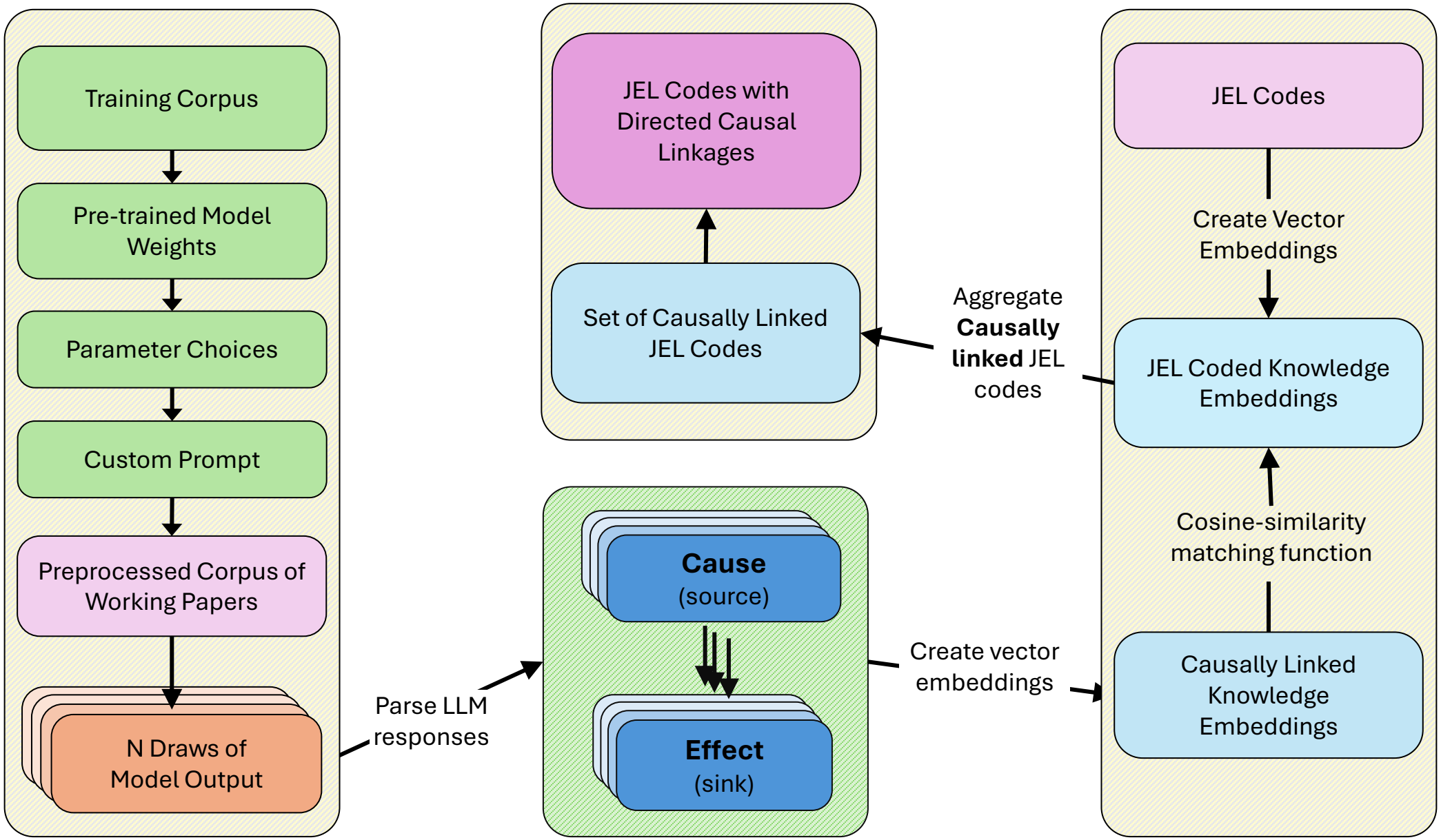
Esther Duflo; Abhijit Banerjee; Rachel Glennerster; Cynthia G. Kinnan



Summary:

Evaluates the impact of introducing microfinance in India, finding increased borrowing and investment but limited effects on consumption and development outcomes.

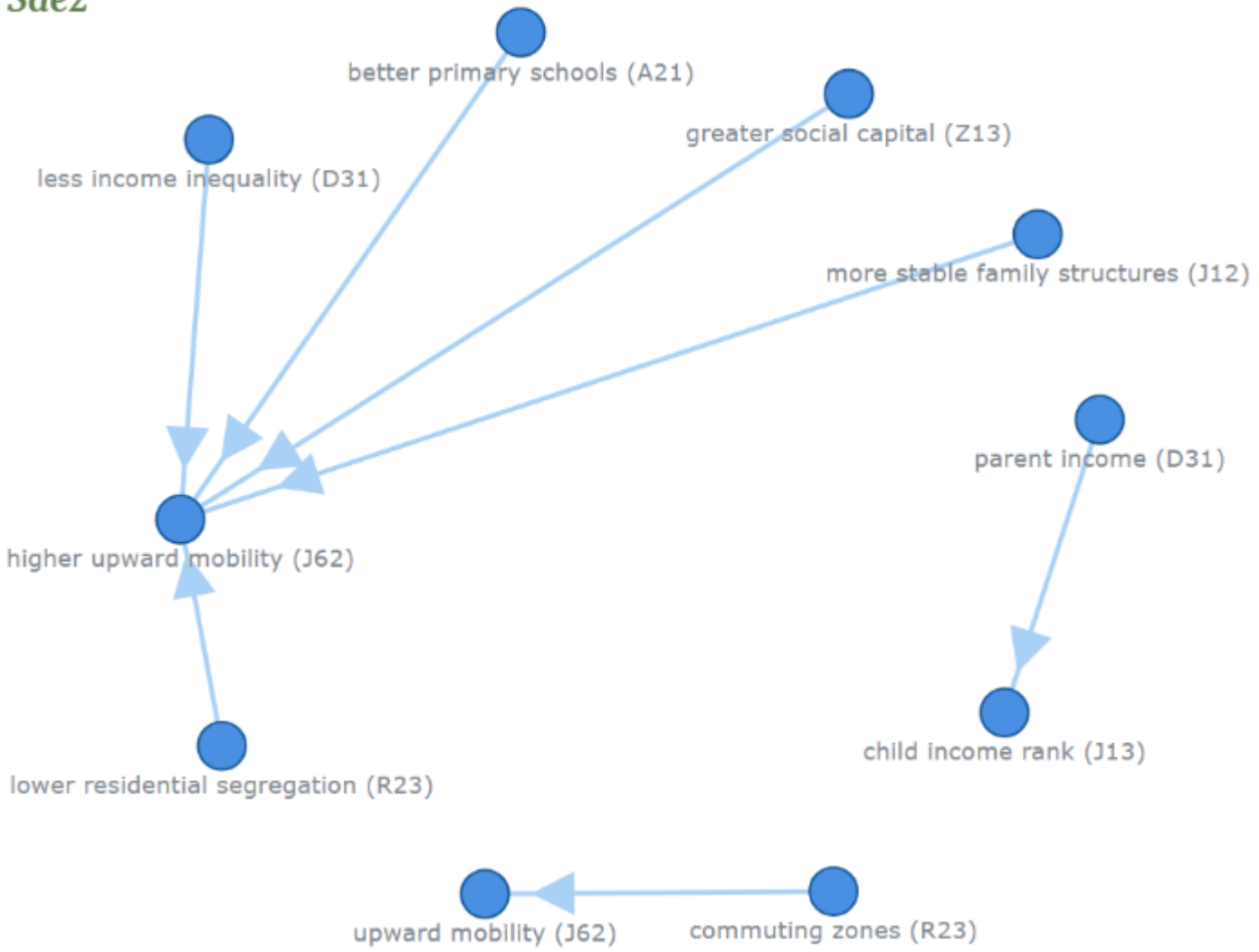
Mapping Causal Linkages Between JEL Codes Using AI



Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States (2014, QJE)

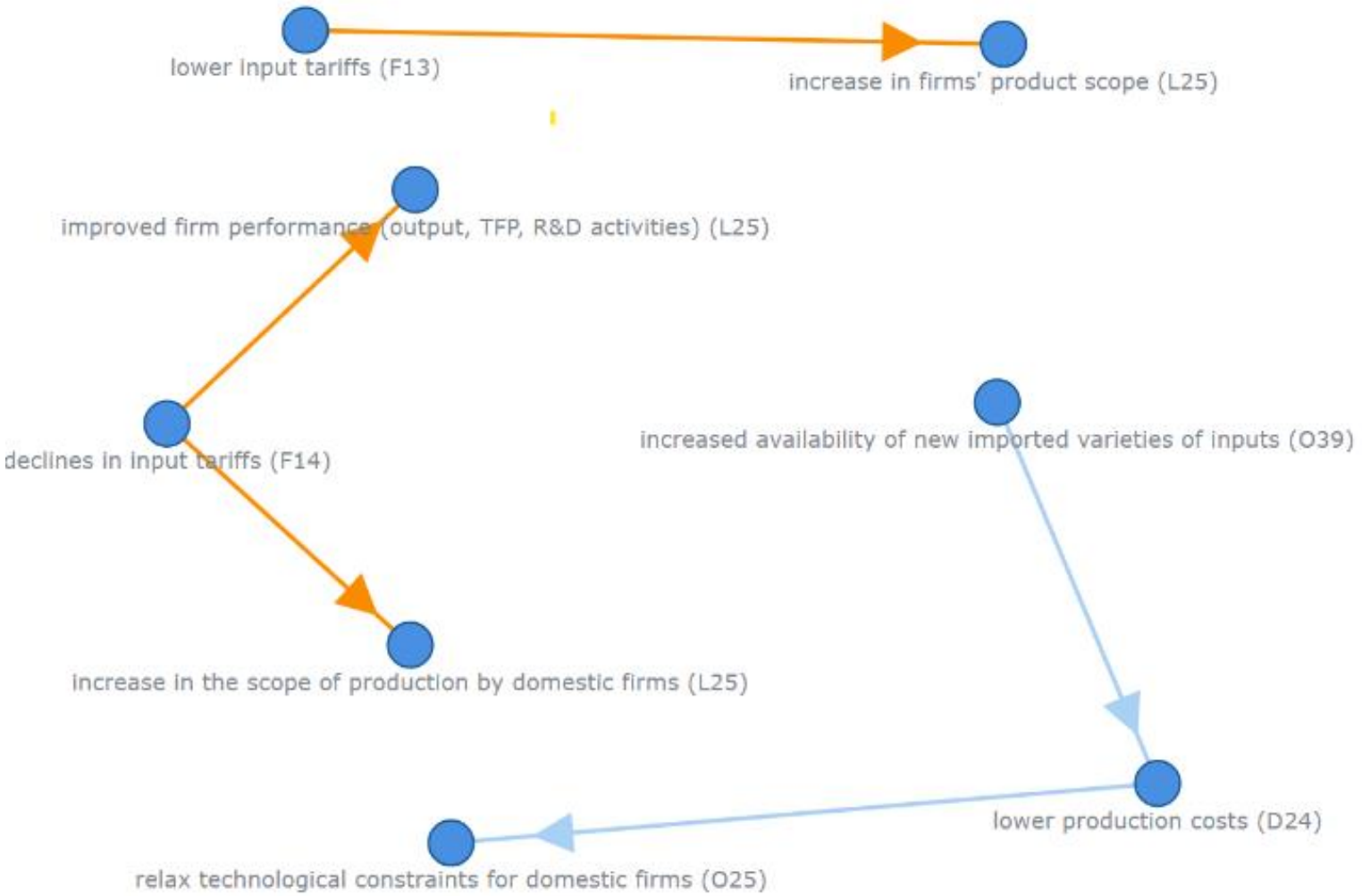
Raj Chetty, Nathaniel Hendren, Patrick Kline, Emmanuel Saez

Summary:
Analyzes U.S. intergenerational income mobility, identifying factors like less segregation and better schools that correlate with higher upward mobility.



Imported Intermediate Inputs and Domestic Product Growth: Evidence from India (2010, QJE)

Pinelopi K. Goldberg, Amit Khandelwal, Nina Pavcnik, Petia Topalova



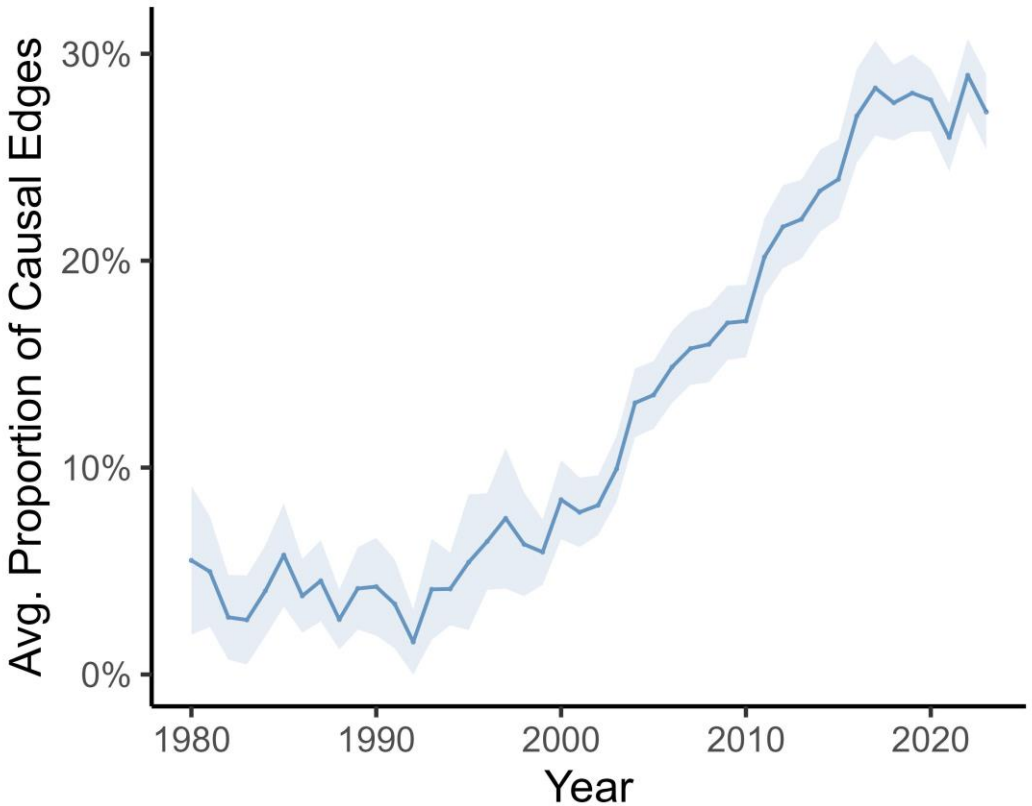
Summary: Examines how access to imported inputs due to lower tariffs boosts Indian firms' product growth and performance by relaxing technological constraints.

Edges that are evidenced by causal inference methods are in orange, and the rest are in light blue.

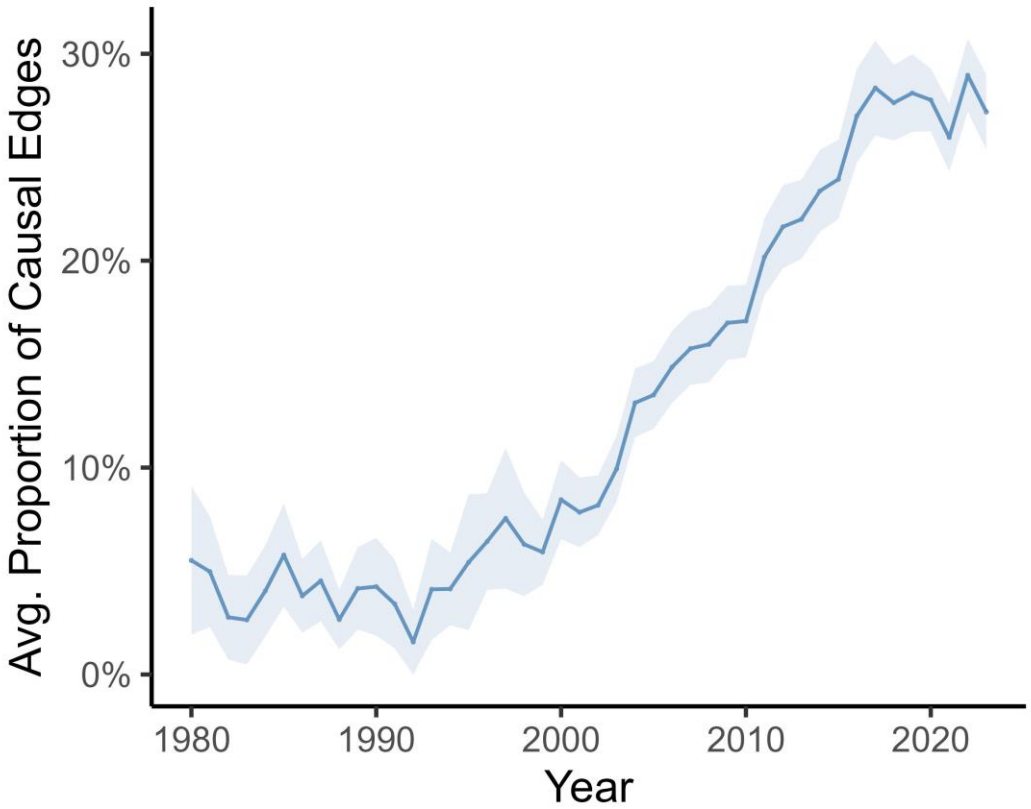
The Causal Graph of Economics Literature

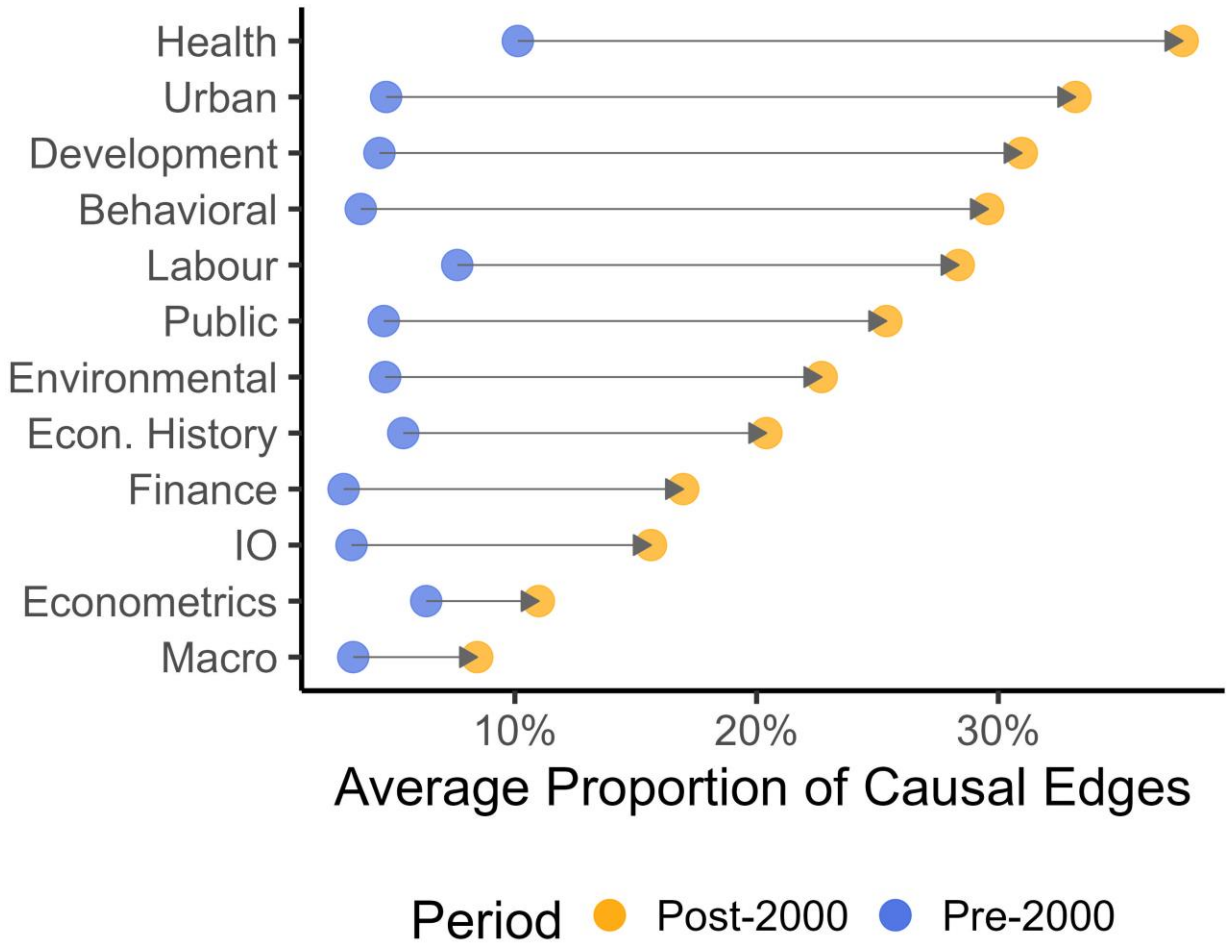
- We construct a knowledge graph for each paper, where nodes represent economic concepts (JEL codes), and edges represent claims from a source node to a sink node.
- Use of JEL codes is primarily for trackability, allowing us to group related concepts (e.g., cost of living, price level increases, inflation, deflation) into one (e.g. E31 - Inflation).
- Claims are classified as **causal** if they are supported by causal inference methods such as Difference-in-Differences (DiD), Instrumental Variables (IV), Randomized Controlled Trials (RCTs), and others.
- We use the **proportion of causal edges in a paper** to measure extent to which economists have increasingly adopted rigorous causal inference methods in their work, indicative of the credibility revolution.

The Causal Graph of Economics



The Causal Graph of Economics





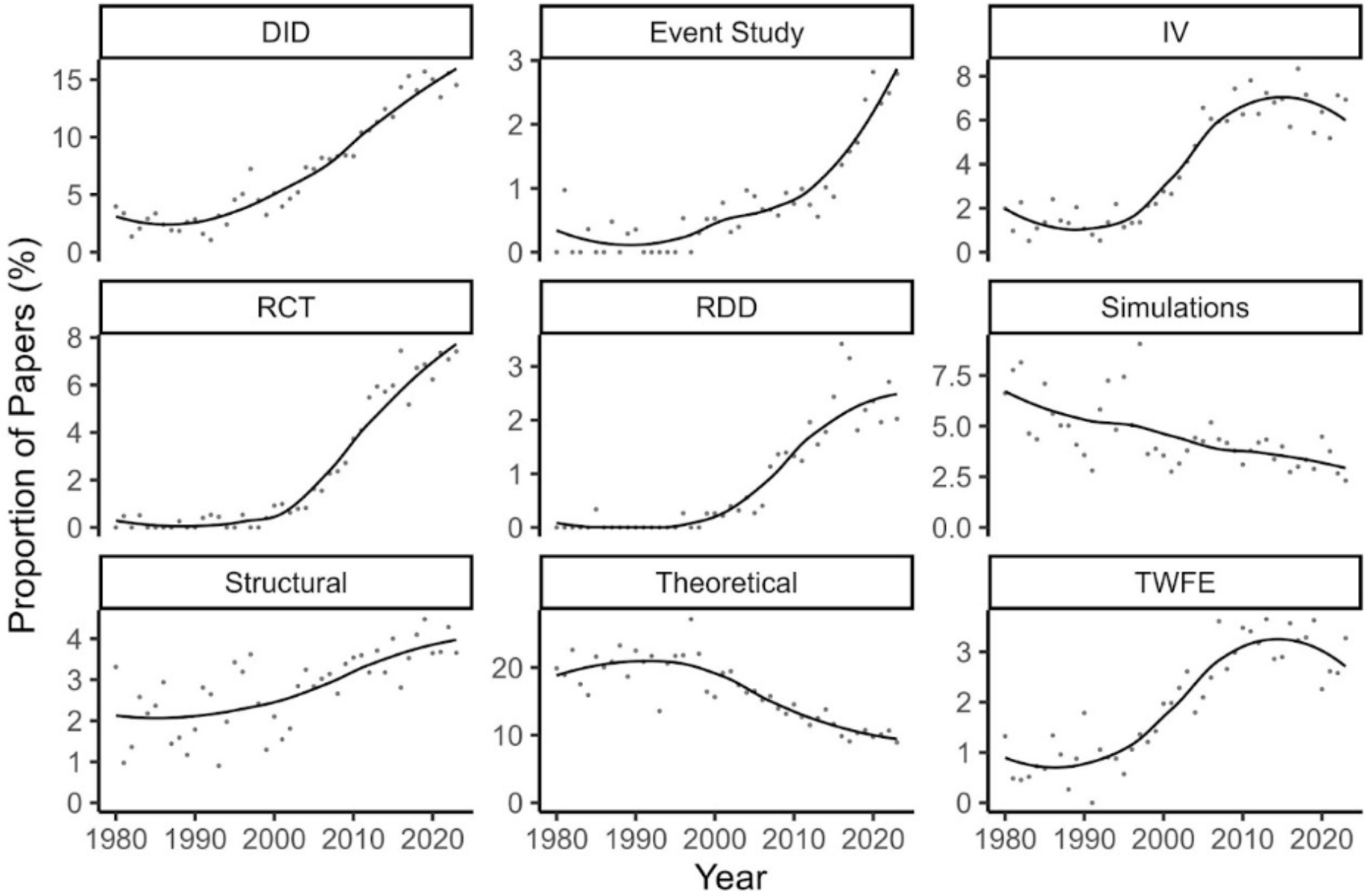
Breakdown by Methods

Growth of "Mostly Harmless Econometrics"

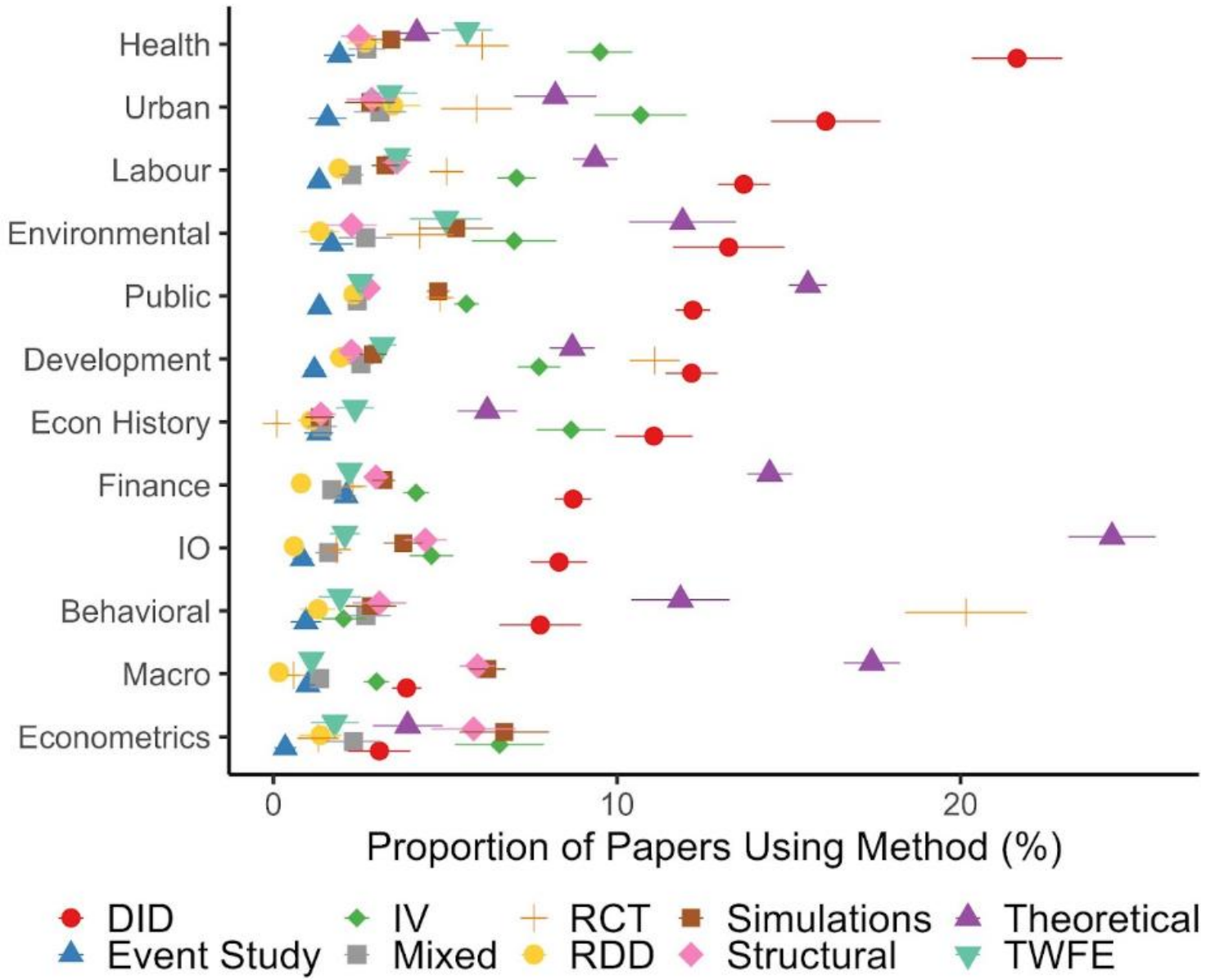
Significant increase in the use of methods like Difference-in-Differences (from ~4% in 1980 to over 15% recently), Instrumental Variables, and Randomized Controlled Trials.

Decline in Theoretical Work

The proportion of theoretical and non-empirical research has declined from approximately 20% in 1980 to under 10% in 2023.



Empirical Methods across Fields



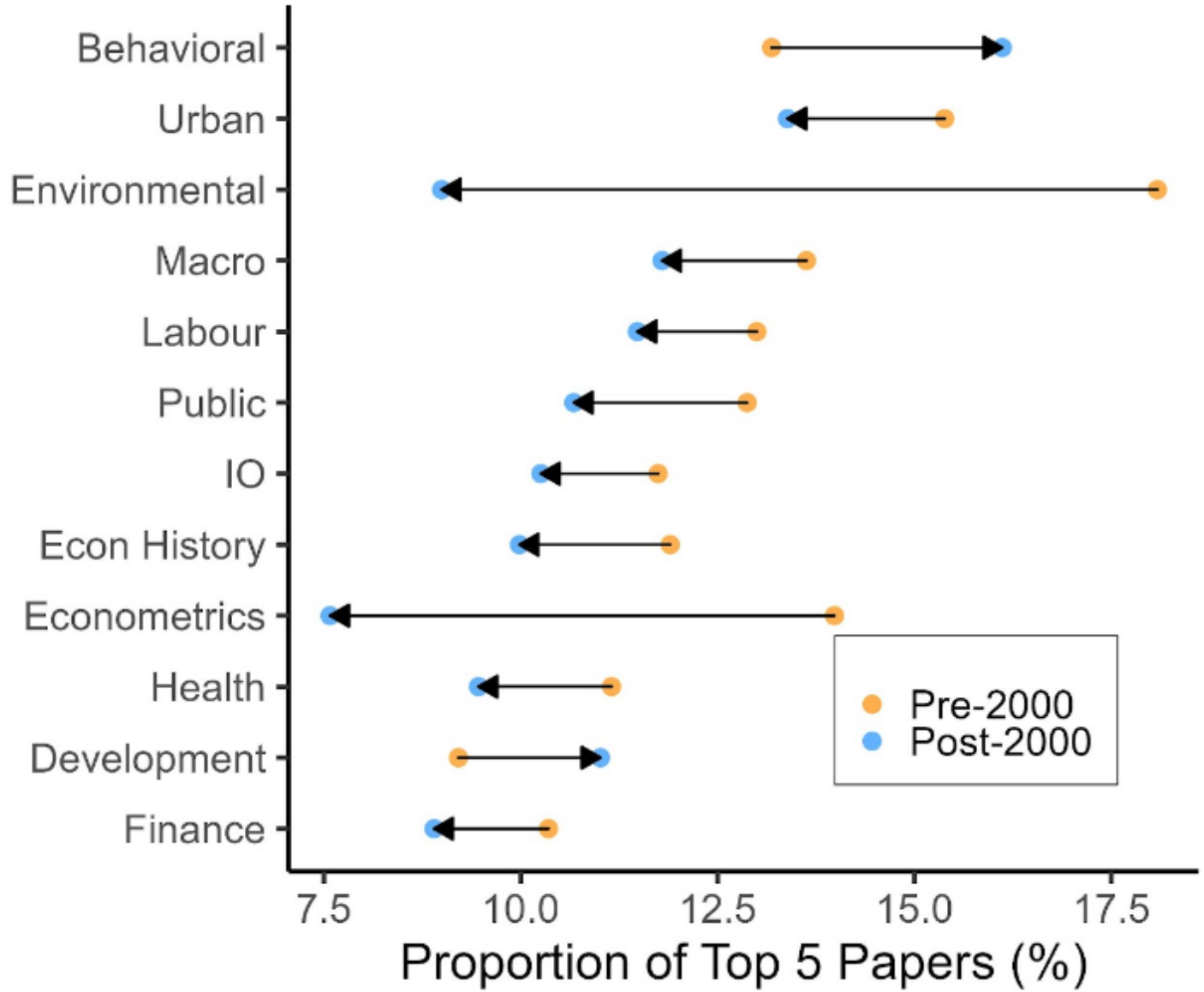
Different fields have adopted methods at varying rates.

- Health, Urban and Labour heavily utilize **Diff-in-Diff**.
- Behavioural and Development Economics prominently feature **RCTs**
- **Theory** is most prevalent in IO and Macro.

Some Context

Types of papers published in top 5 journals

- With fixed supply of top 5, and an increasing demand for publication avenues, the share of papers total papers published in top 5 has reduced.
- This also increases the status of a Top-5
- Some fields have gained disproportionate interest by the Top-5: Behavioural and Development.



Predictors of Publication in Top 5

Top 5 Journals: *AER*, *QJE*, *JPE*, *ReStud*, *ECMA*

- Importance of Top 5 for Career progression

  (Heckman and Moktan, 2020; Card and DellaVigna, 2013)

- Evolution in research interests, methodological innovations, journal policies

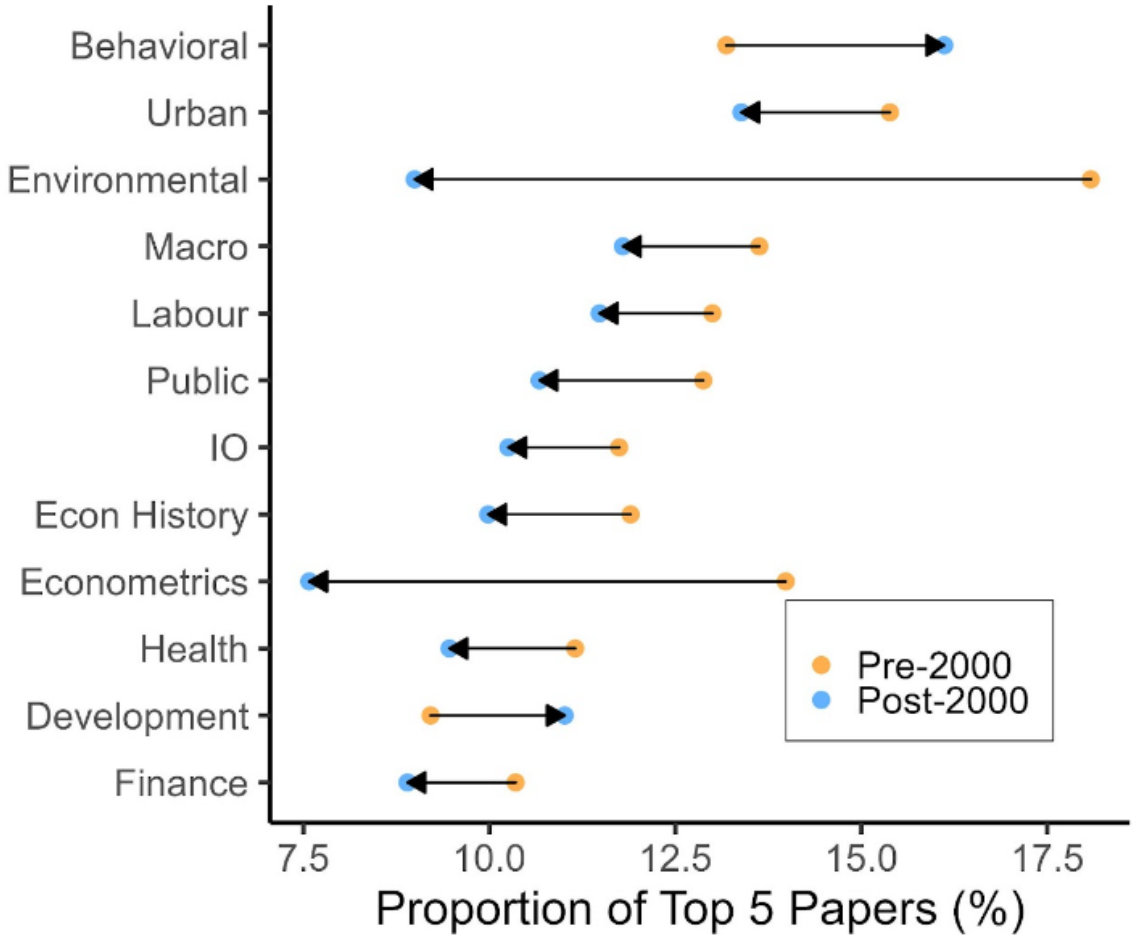
Predictors of Publication in Top 5

Top 5 Journals: *AER, QJE, JPE, ReStud, ECMA*

- Importance of Top 5 for Career progression

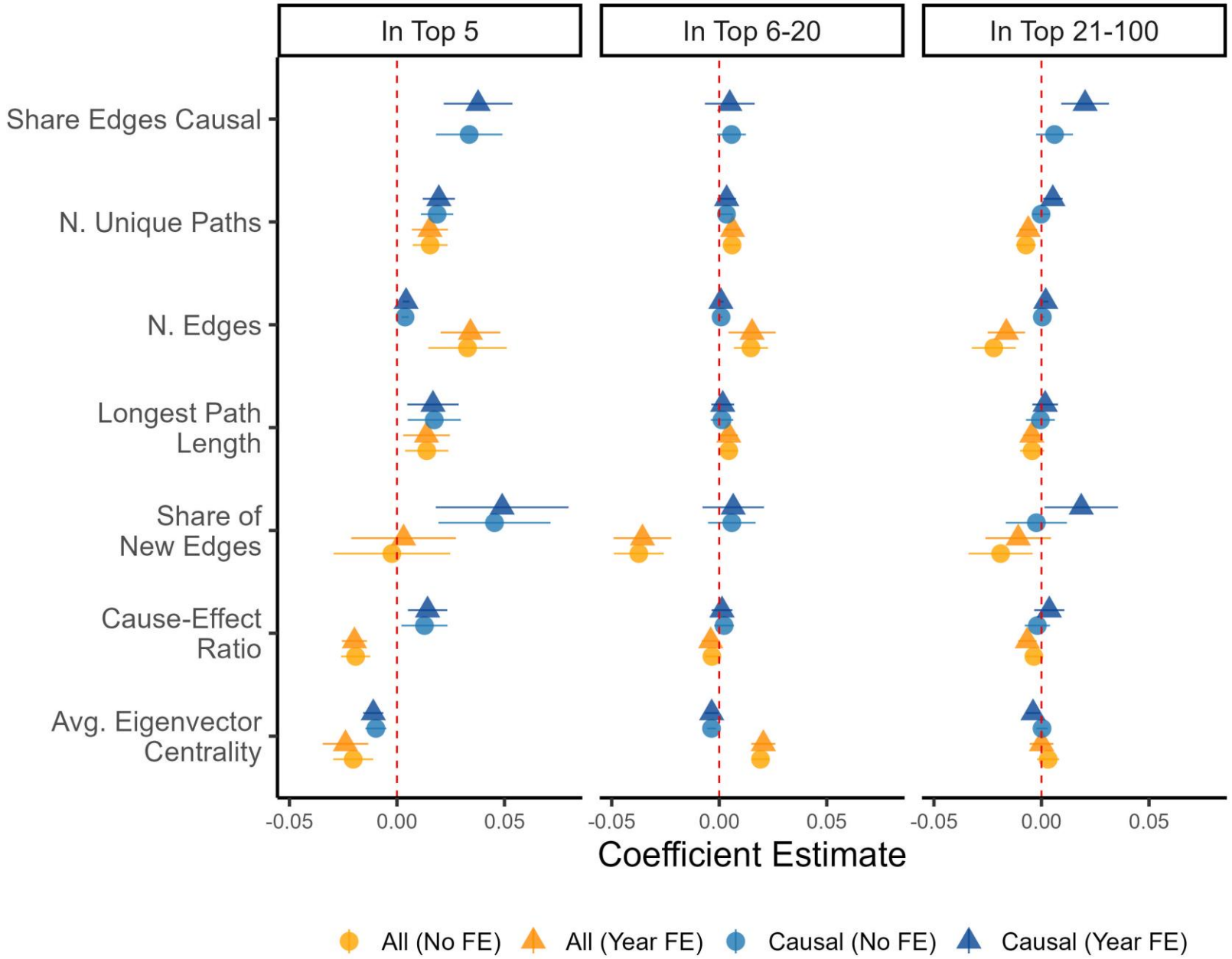
  (Heckman and Moktan, 2020; Card and DellaVigna, 2013)

- Evolution in research interests, methodological innovations, journal policies



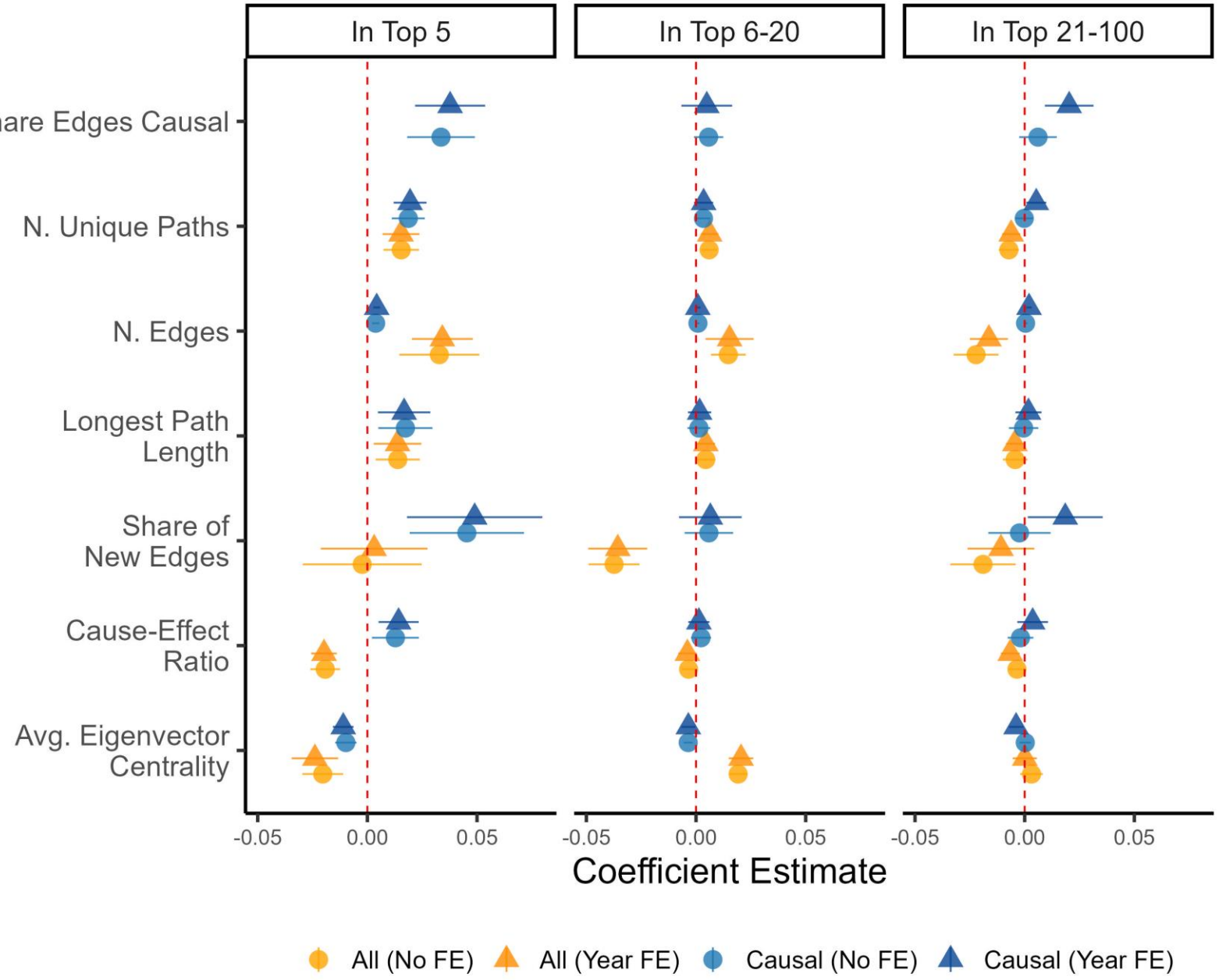
- With fixed supply of top 5, and an increasing demand for publication avenues, the share of papers total papers published in top 5 has reduced.
- This also increases the status of a Top-5
- Some fields have gained disproportionate interest by the Top-5: Behavioural and Development.

What the journals are more likely to Publish



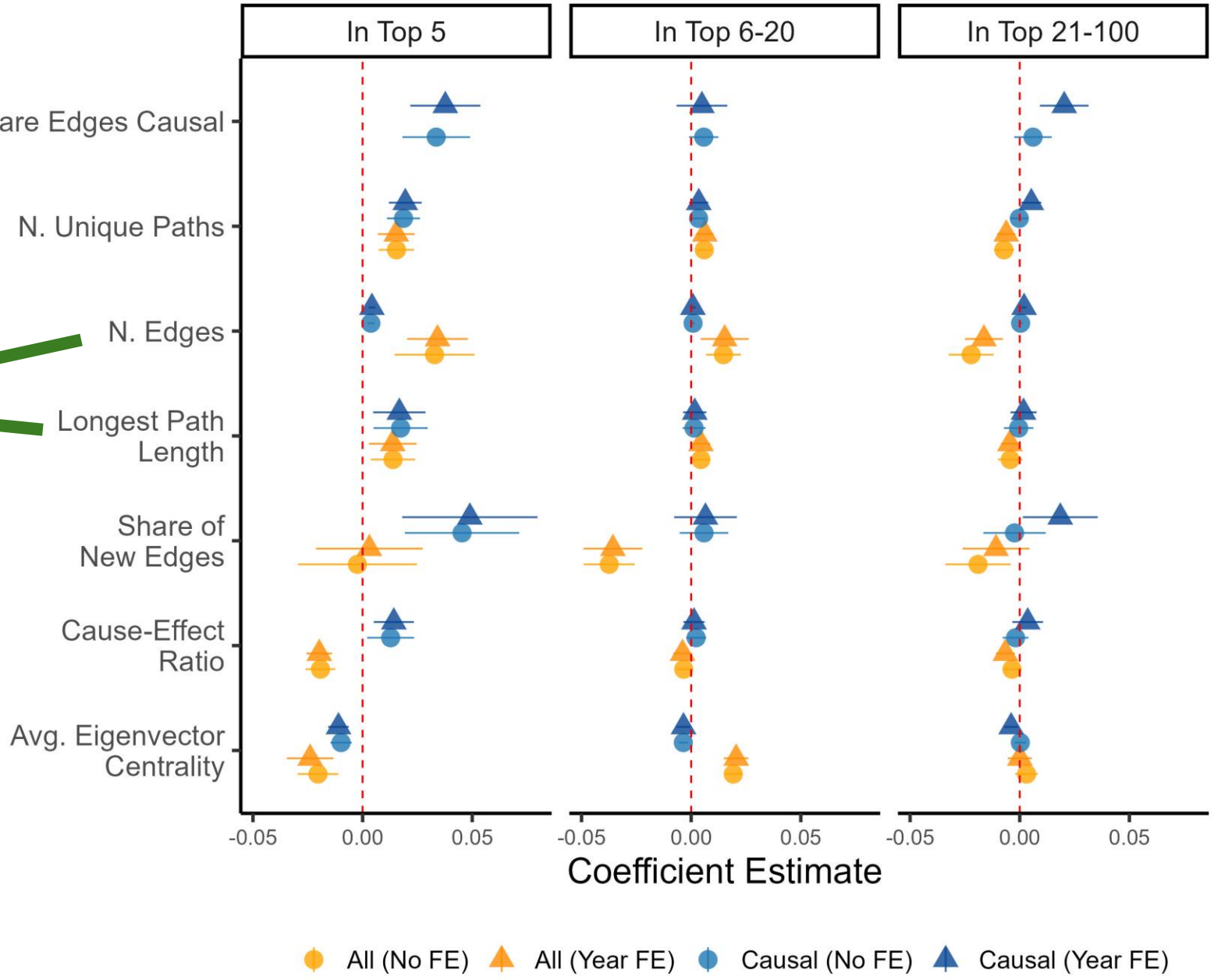
What the journals are more likely to Publish

- Higher proportion of causal edges increases likelihood of publication in Top 5 journals.



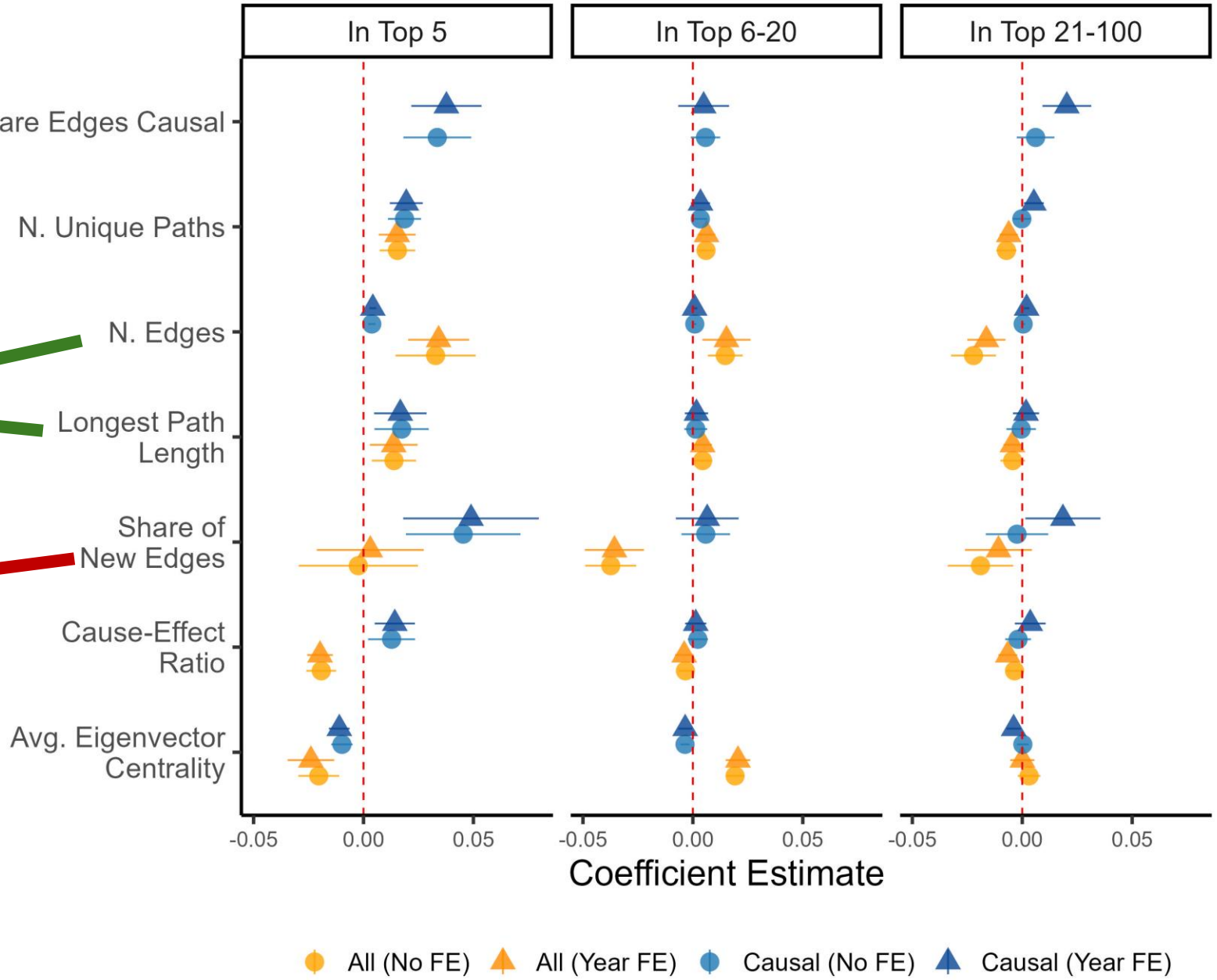
What the journals are more likely to Publish

- Higher proportion of causal edges increases likelihood of publication in Top 5 journals.
- Narrative complexity (unique paths, longest path length) positively associated with top journal publications.



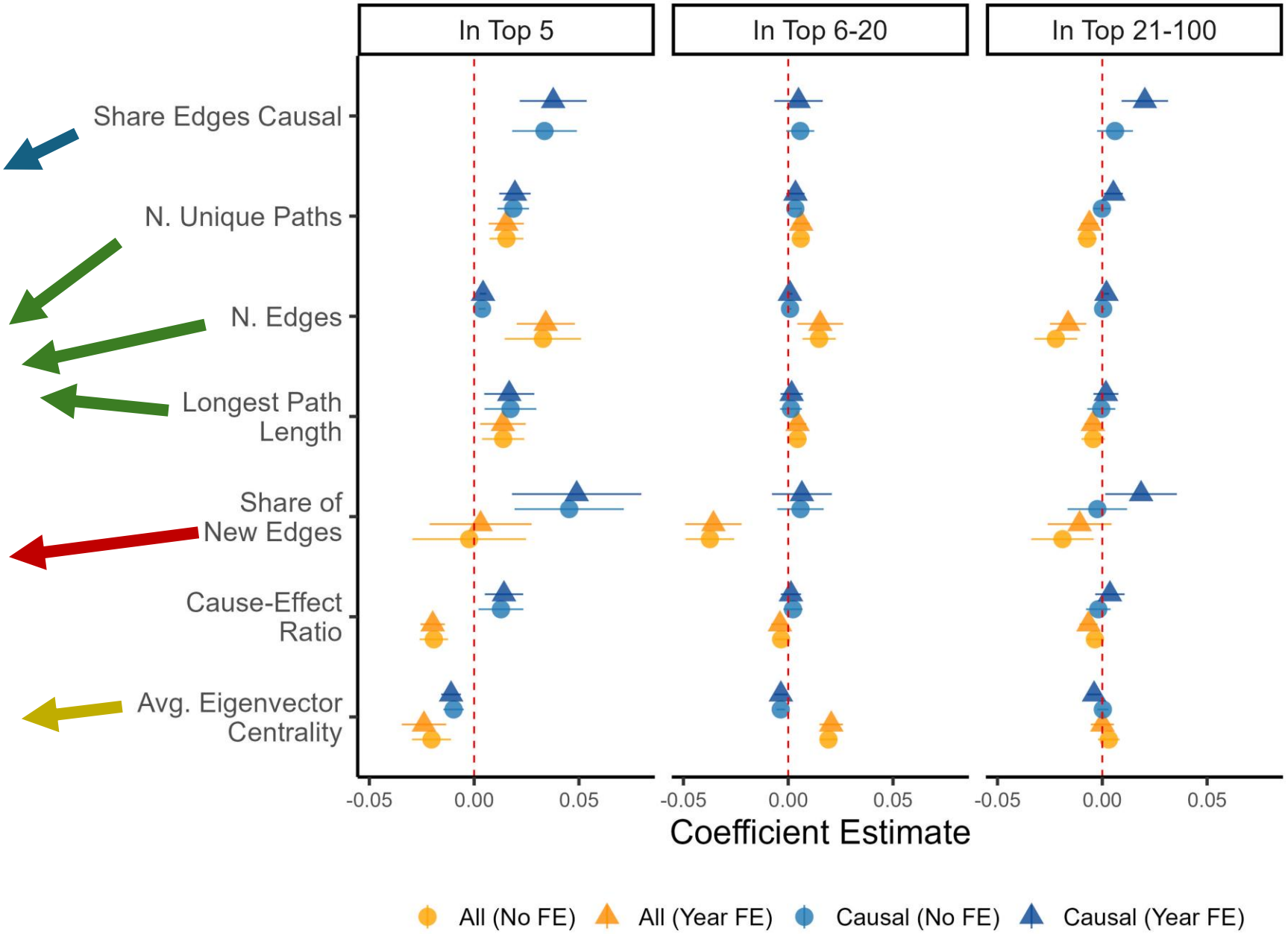
What the journals are more likely to Publish

- Higher proportion of causal edges increases likelihood of publication in Top 5 journals.
- Narrative complexity (unique paths, longest path length) positively associated with top journal publications.
- Novel causal relationships more likely to be published in Top 5 journals.

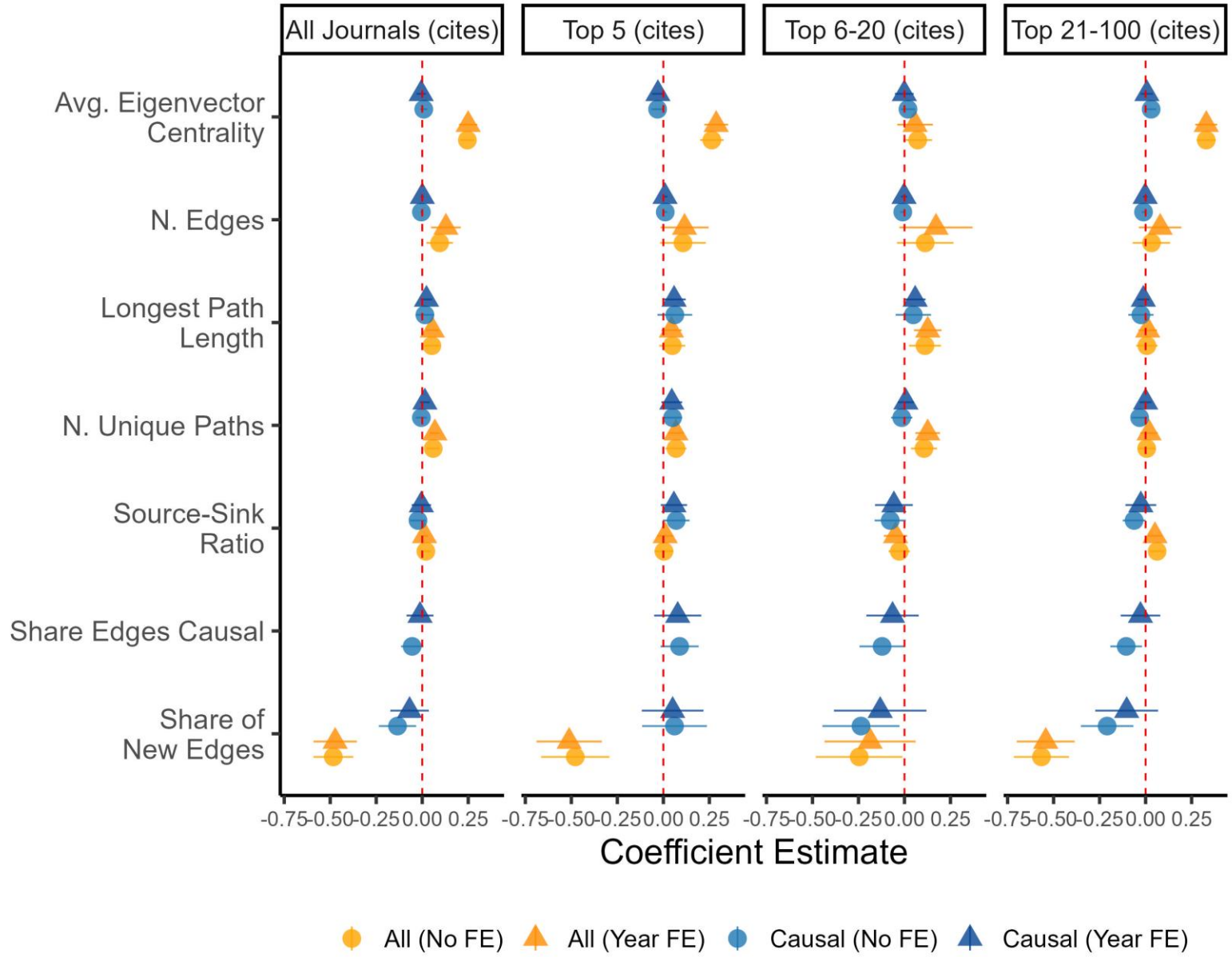


What the journals are more likely to Publish

- Higher proportion of causal edges increases likelihood of publication in Top 5 journals.
- Narrative complexity (unique paths, longest path length) positively associated with top journal publications.
- Novel causal relationships more likely to be published in Top 5 journals.
- Engagement with less central concepts increases chances in Top 5 journals.

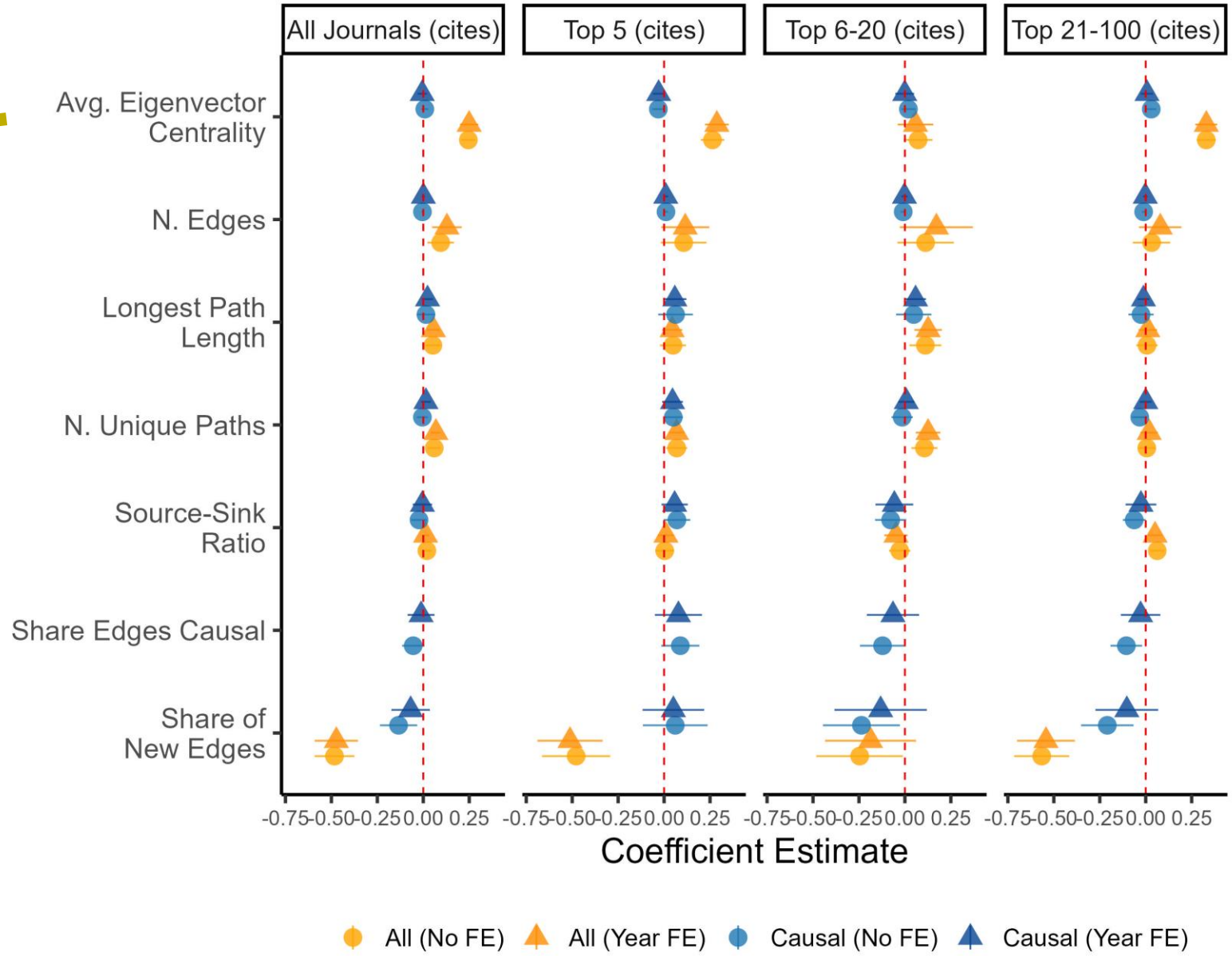


Predictors of Citation Counts, given publication



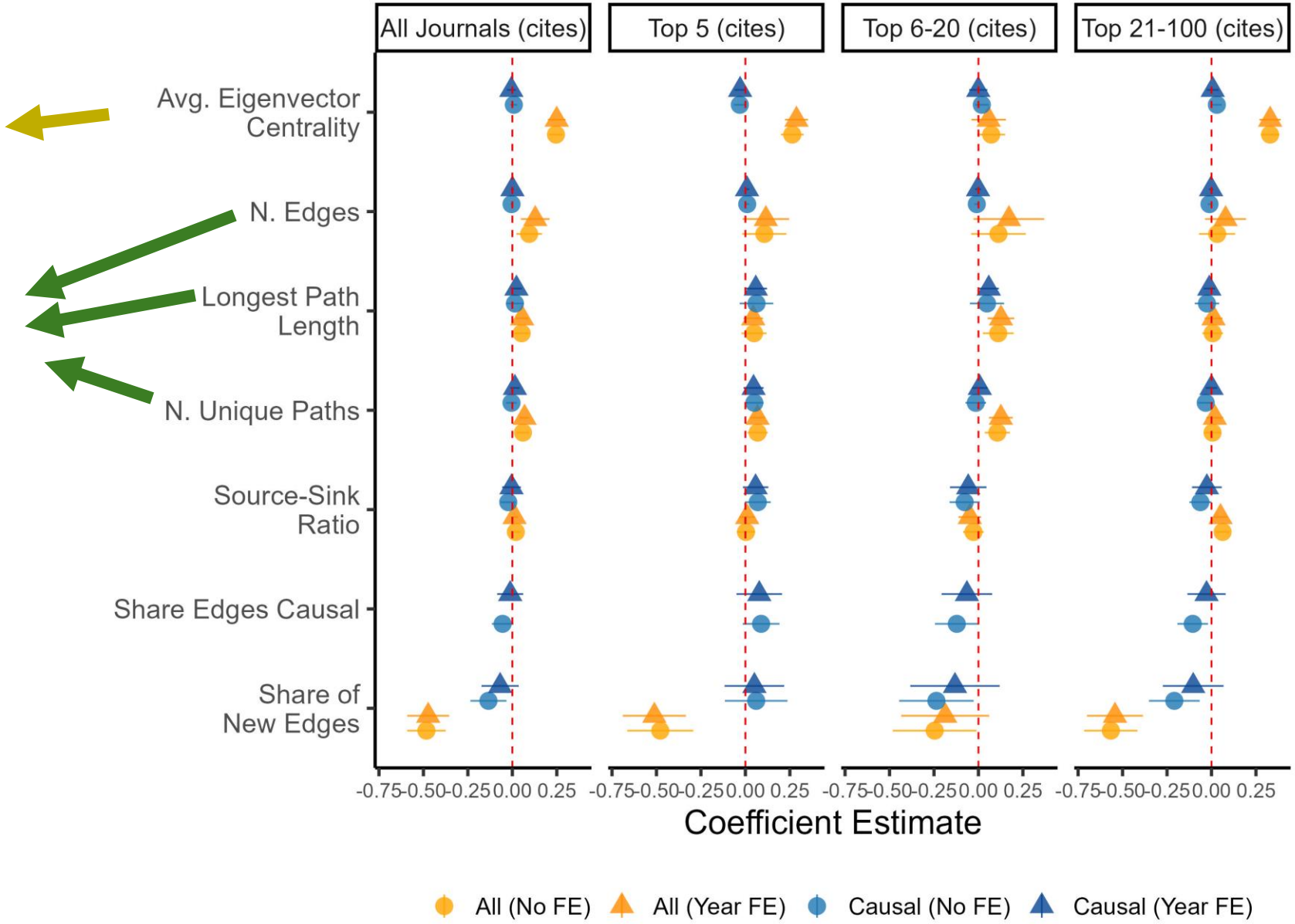
Predictors of Citation Counts, given publication

- Papers focusing on central concepts receive more citations once published.



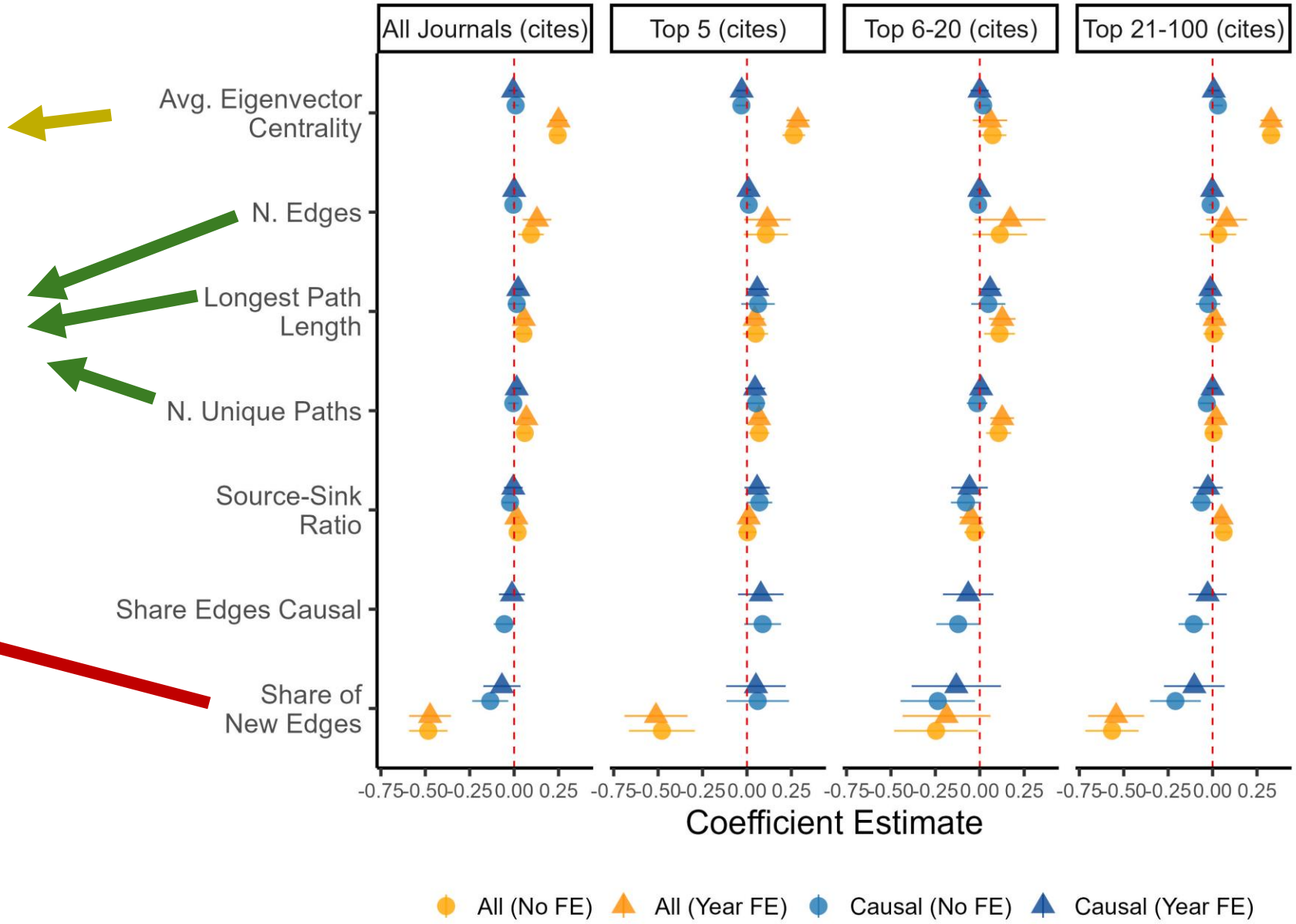
Predictors of Citation Counts, given publication

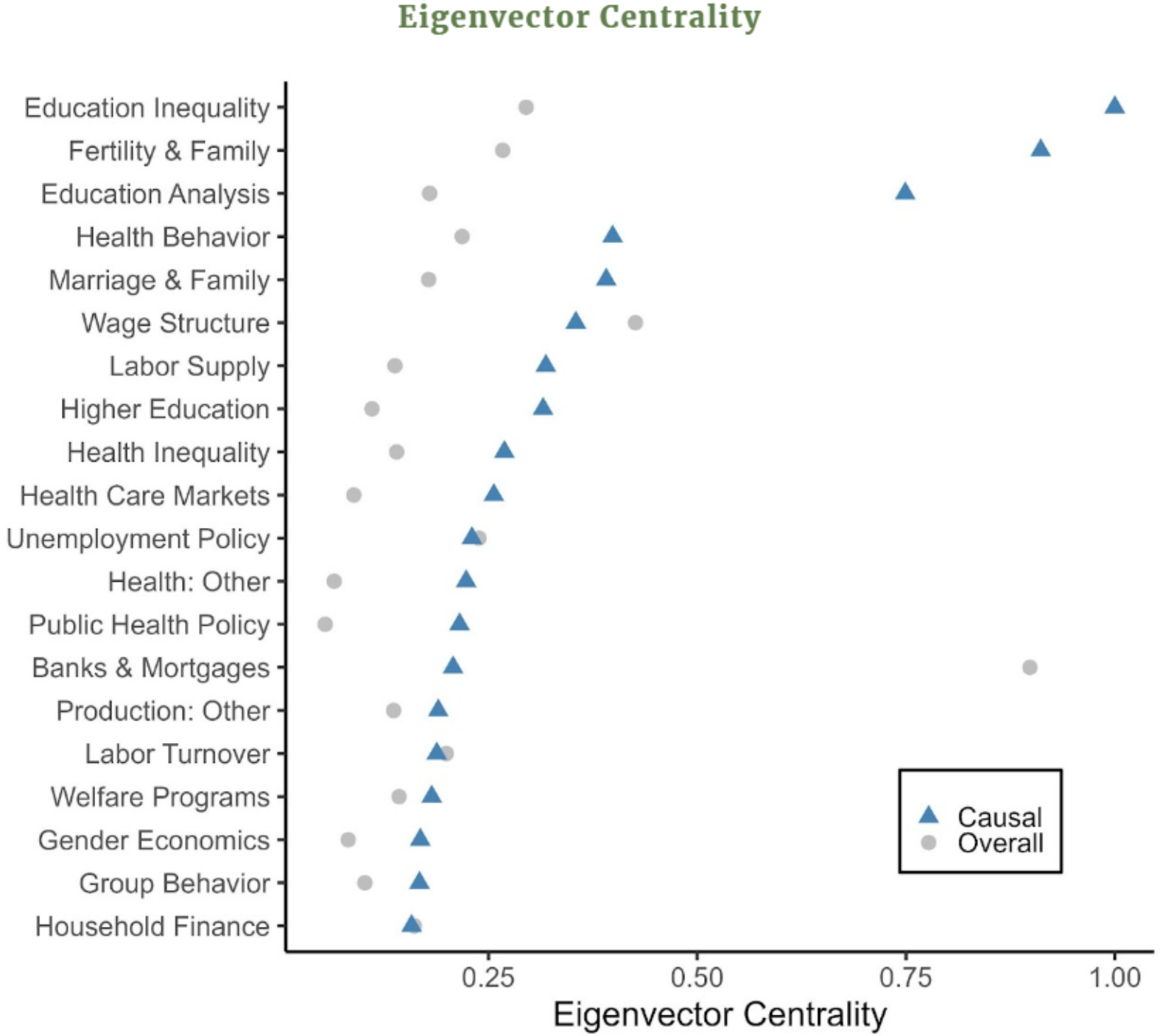
- Papers focusing on central concepts receive more citations once published.
- Narrative complexity *only slightly* enhances citation counts, especially in Top 5 and Top 6–20 journals.



Predictors of Citation Counts, given publication

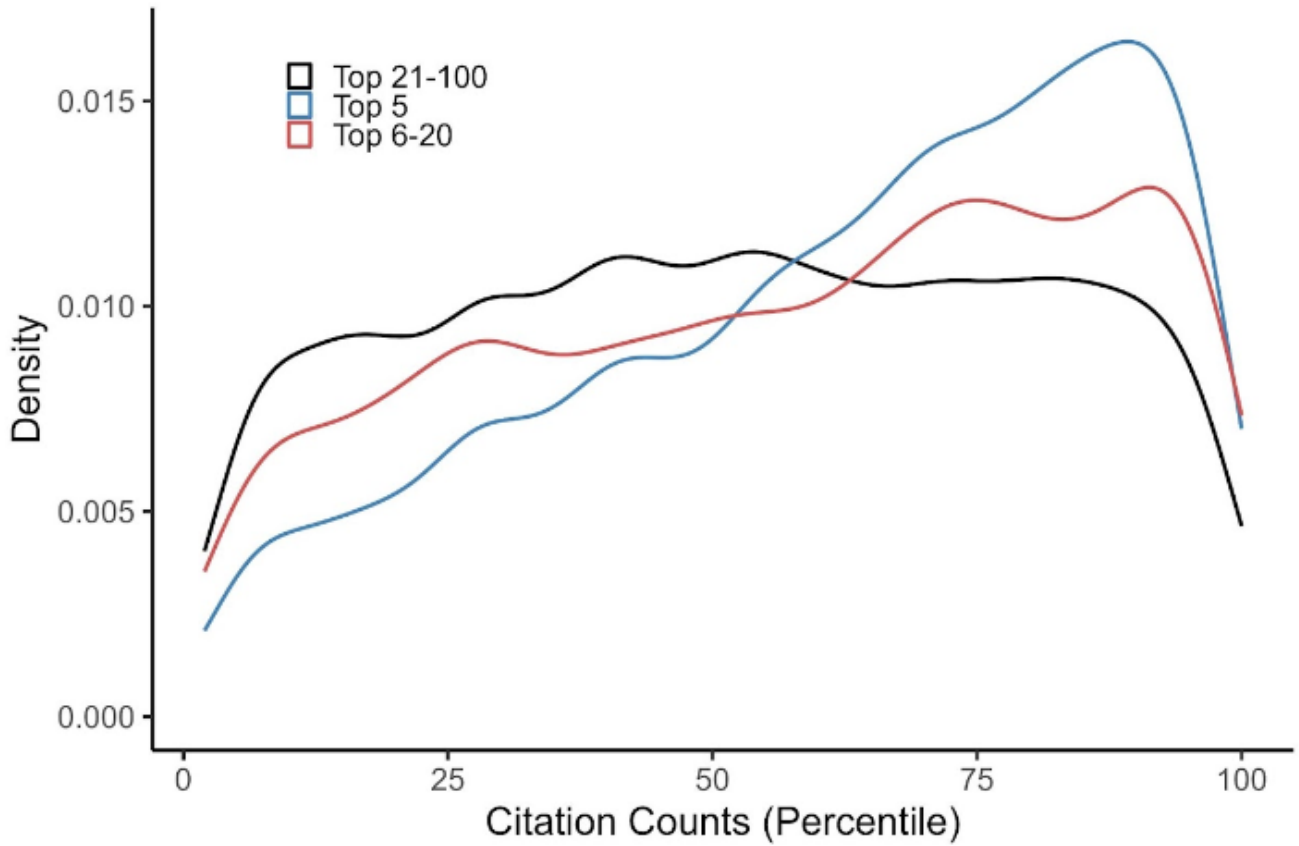
- Papers focusing on central concepts receive more citations once published.
- Narrative complexity *only slightly* enhances citation counts, especially in Top 5 and Top 6–20 journals.
- Novelty in causal relationships does not necessarily increase citations.





- **Concept Influence Measured by Centrality**: Average eigenvector centrality quantifies how influential economic concepts are within the knowledge graph.
- **Top Concepts Differ in Causal Graph**: In the *overall* graph, the most central nodes are G21 (Banks and Mortgages), J31 (Wage Structure), and I24 (Education and Inequality). In the *causal* graph, they shift to I24 (Education and Inequality), J13 (Fertility and Family), and I21 (Analysis of Education).
- **Causal Research Focuses on Specific Fields**: Causal inference methods are predominantly applied in education, family, and health economics, indicating these areas are more amenable to causal analysis.
- **Skewed Centrality Distribution**: Centrality scores are highly skewed, with few nodes having very high centrality—especially in the causal graph (e.g., I24 normalized to 1.0).
- **Implications for Research Focus**: Papers involving highly central nodes in the causal graph contribute to established research areas using rigorous methods, while those with less central nodes explore novel or specialized topics.

Distribution of Citation Percentiles by Journal Category



- This figure displays kernel density plots of the citation percentiles for papers published in Top 5, Top 6-20, and Top 21-100 journals.
- The plot shows that while papers published in higher-ranked journals tend to receive more citations on average, the most highly cited papers are more evenly distributed across journal categories.
- This suggests that exceptionally influential papers can emerge from a wide range of journals.

Risks to Replication of Credibility Revolution

While the advanced causal inference methods increases credibility of causal results, there are two concerns that make studies hard to replicate:

- decline in reporting of null results
- increase in use of private sector data

Risks to Replication of Credibility Revolution

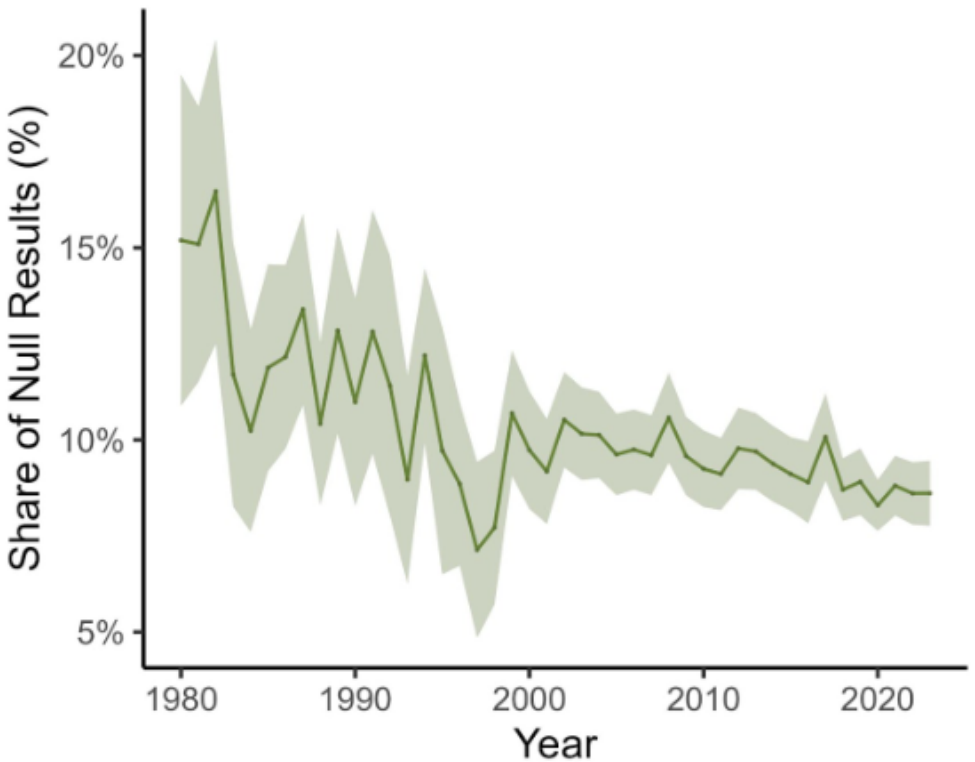
While the advanced causal inference methods increases credibility of causal results, there are two concerns that make studies hard to replicate:

- decline in reporting of null results
- increase in use of private sector data

1. Decline in Reporting of Null Results

The Importance of Transparency in Research Findings

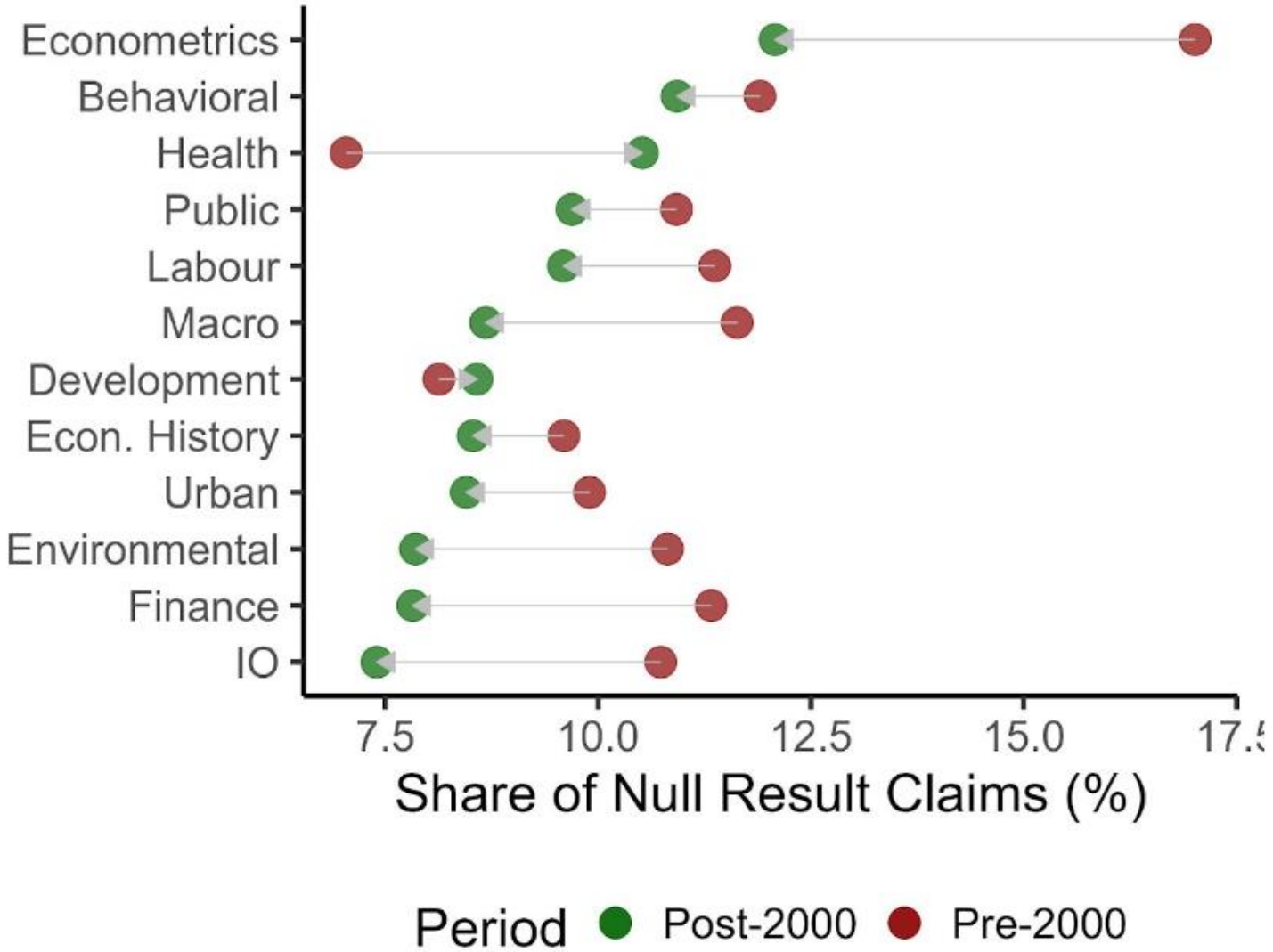
- The reporting of null results is critical for scientific transparency and the accumulation of knowledge.
- Null results provide valuable information about the absence of expected effects, helping to prevent publication bias and ensuring a more accurate understanding of economic phenomena (Rosenthal 1979, Sterling 1959).
- Despite their importance, null results are often underreported due to the perceived lower likelihood of publication and the undervaluation of such findings in the academic community (Brodeur et al. 2016, Chopra et al. 2024).



Risks to Replication of Credibility Revolution

A reflection of research practices and publication norms across field

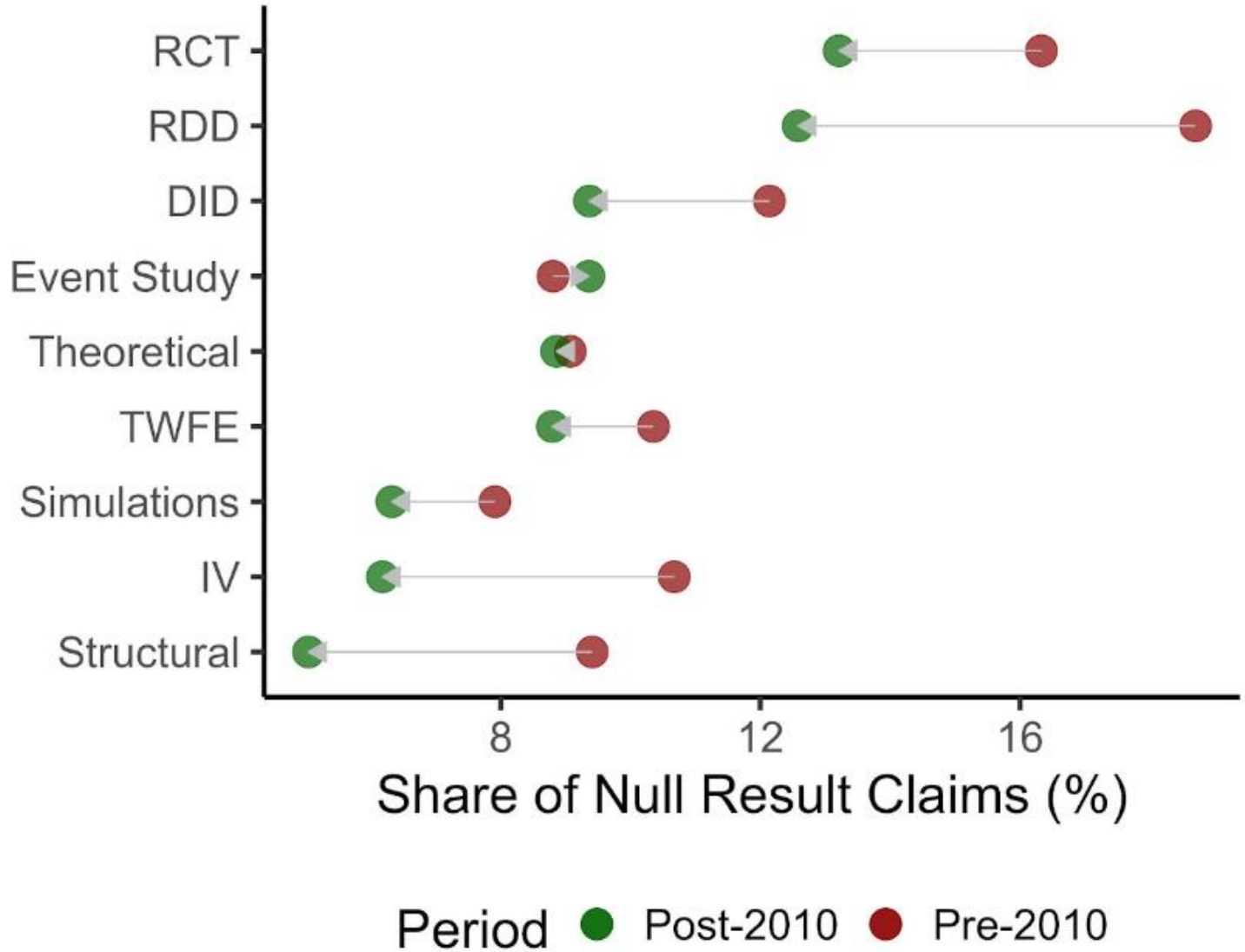
- We observe a general decrease across most fields in the post-2000 period.
- Highest levels: Fields such as Econometrics and Behavioural Economics report higher shares of null results.
- Health Economics shifted from least to one of the highest reporter of Nulls.
- Finance and IO report lowest shares post-2020



Risks to Replication of Credibility Revolution

Large variations across methods

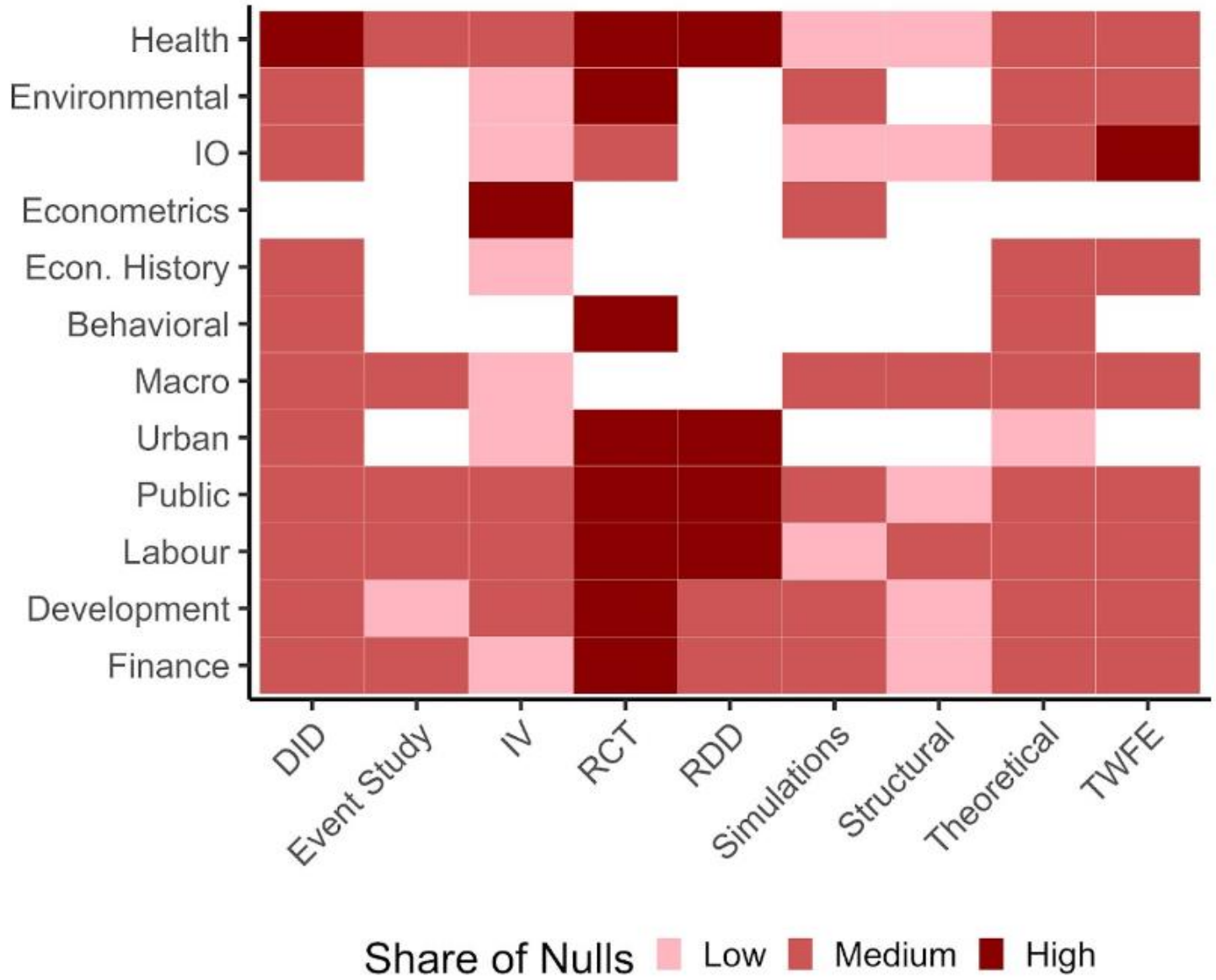
- Causal inference design methods like RCTs, DiD, and RDD are associated with higher shares of null results.
- Methods like Structural Estimation, IV and Simulations see the lowest share of Null result claims.
- These patterns may reflect the differing nature of these methods.
- Experimental and quasi-experimental methods like RCTs and RDDs, designed for rigorous causal inference, may often result in null findings when interventions do not produce significant effects.
- Transparent reporting of such results is important to avoid publication bias.



Risks to Replication of Credibility Revolution

Large variations across methods

- RCTs and RDDs have *relatively* high null results across all fields.
- Opposite is true for IV, with the exception of papers that are primarily about econometrics.

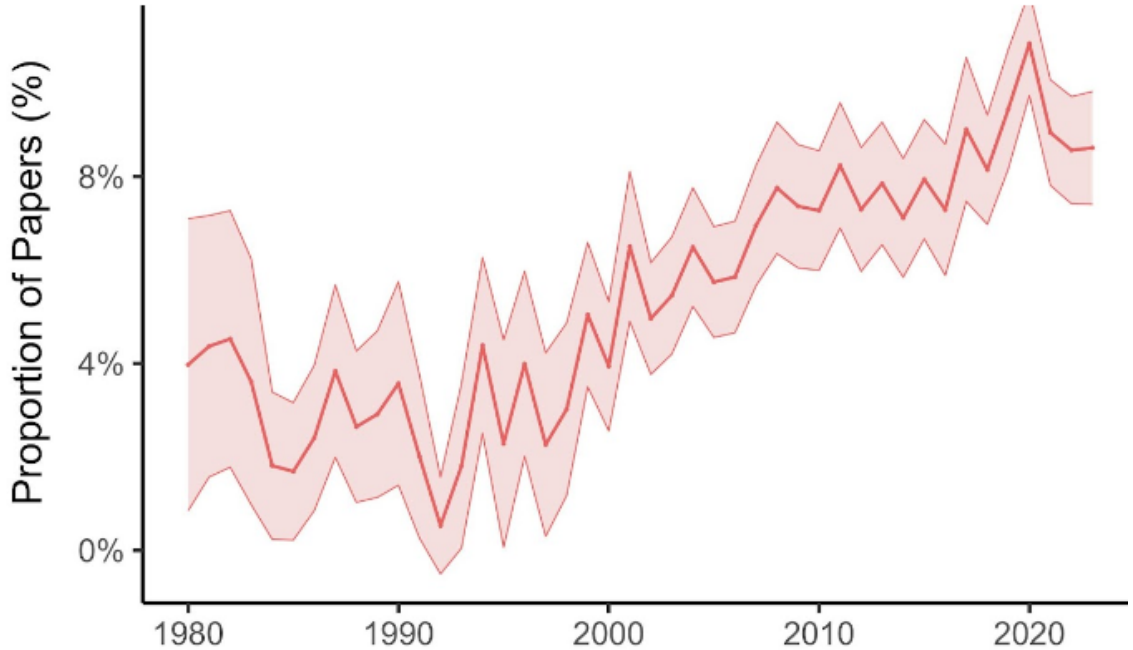


Risks to Replication of Credibility Revolution

2. Use of Private Data in Economics Research

Implications for Data Accessibility and Replicability

- The use of proprietary data exacerbates the problem, limiting other researchers' ability to replicate studies or test alternative hypotheses.
- Open data is not widely practiced in economics (Andreoli-Versbach & Mueller-Langer 2014).
- The need for policies that balance privacy concerns with the benefits of data accessibility for scientific advancement (Fetzer 2022).
- Data privacy regulations like the GDPR have introduced additional barriers to data sharing.
- In response to these challenges, Miguel (2021) documents the adoption of open science practices in economics, such as pre-registration and data sharing, noting a rapid transition toward increased transparency.



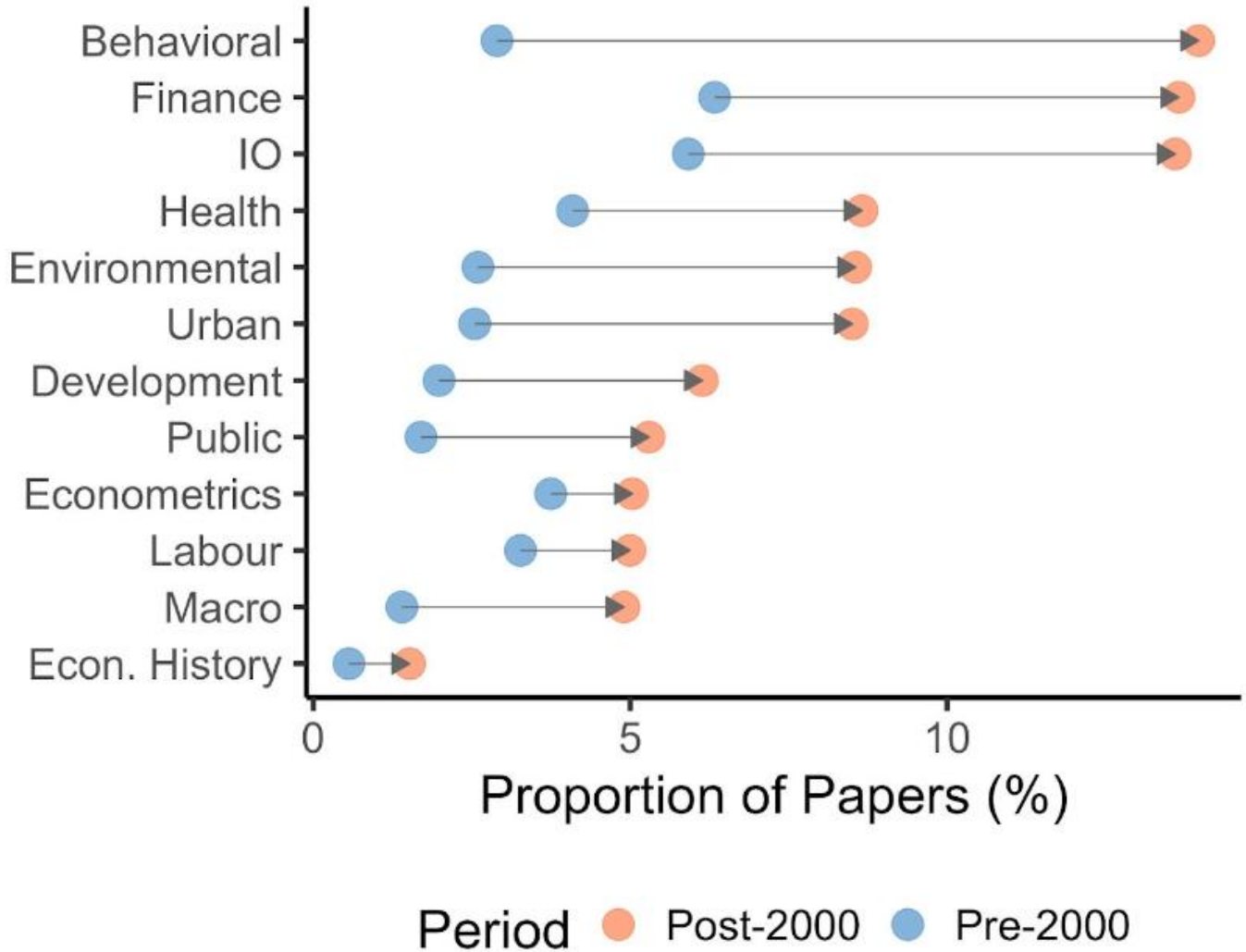
Increase Over Time: Use of private or proprietary data rose from about 4% in 1980 to approximately 8.6% in 2023.

This trend reflects the greater availability of granular, individual-level data collected by private companies, as well as increased collaboration between researchers and private entities.

Risks to Replication of Credibility Revolution

A reflection of research practices and publication norms across field

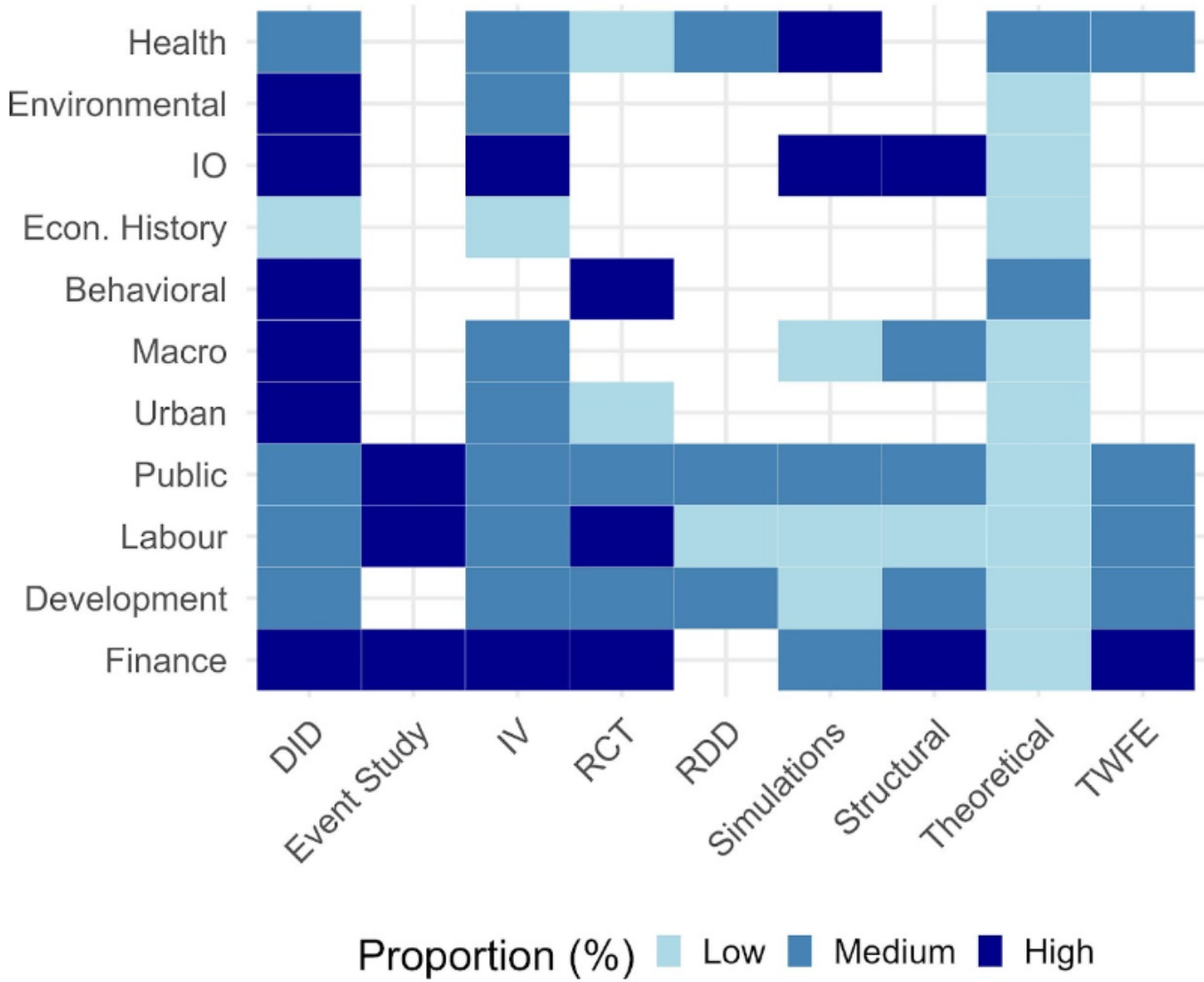
- Fields such as Finance, IO, and Behavioral Economics exhibit higher proportions of papers using private data in the post-2000 period.
- These increases may reflect the nature of research in these fields, which often relies on firm-level or experimental data that is not publicly available.
- The use of proprietary datasets allows researchers to conduct detailed analyses of financial markets, consumer behaviour, and firm dynamics.



Risks to Replication of Credibility Revolution

Field x Method

- We observe that certain combinations of fields and methods are associated with higher proportions of private data usage.
- For example, DiD in Behavioral Economics (29%) and Finance (20%), and Structural in IO (28%), are associated with higher private data usage.



Conclusion

Study Contributions:

- Comprehensive analysis of over 44,000 working papers.
- Detailed insights into the evolving landscape of empirical methods, causal claims, and data practices in economics.

Recommendations:

- Emphasize transparency, methodological rigor, and the reporting of null results.
- Address challenges related to private data usage and replicability.

Key Findings of our paper

Evolution of Empirical Methods

•**Significant Increase in Causal Claims:** The average proportion of causal claims in papers rose significantly from approximately 5% in 1990 to around 28% in 2020, reflecting the impact of the credibility revolution in economics.

•**Growth in Causal Inference Methods and decline in Theoretical and Simulation Methods**

Complex Narratives Linked to Top Publications:

•**Intricate Causal Narratives Enhance Publication and Citation Impact:** Papers featuring intricate and interconnected causal narratives are more likely to be published in top-tier journals, particularly the top 5 journals, and receive more citations, especially within those journals.

•**Key Measures of Causal Narrative Complexity:** Increases in the number of unique paths and the longest path length in causal knowledge graphs are positively associated with both publication in leading economics journals and higher citation counts. This highlights the value placed on depth and complexity in causal narratives.

•**Depth Over Quantity in Causal Claims:** While the overall number of claims made is positively correlated with top journal publications, the number of causal edges alone does not show the same positive association with publication outcomes or citation counts. This suggests that depth over breadth in causal claims is valued

Key Findings of our paper

Novelty Premium? Novel Causal Relationships Enhance Publication but Not Citation Impact: Papers introducing novel causal relationships that have not been previously documented are more likely to be published in top 5 journals, indicating a premium on originality for publication success. However, this does not necessarily translate into higher citation counts once published.

Central vs. Peripheral Concepts

•**Specialized Topics Gain Recognition in Top Journals, but Central Topics Receive More Citations:** The average eigenvector centrality of nodes is negatively associated with publication in top 5 journals, suggesting that papers engaging with less central, more specialized concepts are more likely to appear in the most prestigious journals. However, once published, papers focusing on more central concepts tend to receive more citations, including in top journals. This indicates a divergence between factors that enhance publication success and those that drive academic influence.

Balance Between Source and Sink Nodes

•**Top Journals Tend to Publish Papers with Multiple Causes Leading to Few Effects in Causal Claims:** In the causal subgraph, top journals publish papers exploring multiple causal factors leading to fewer outcomes (many sources to few sinks), and such papers receive more citations.

•**Opposite Pattern for Non-Causal Relationships**

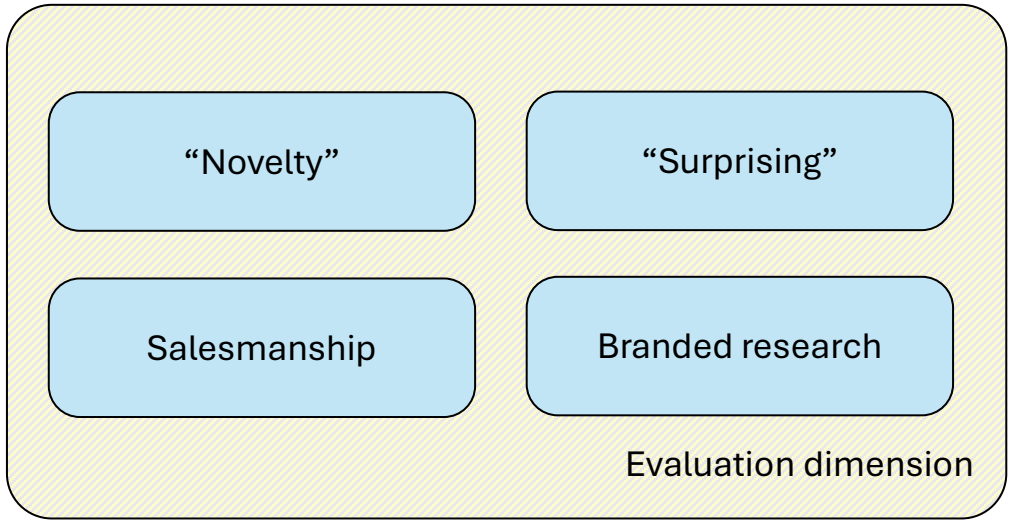
Key Findings of our paper

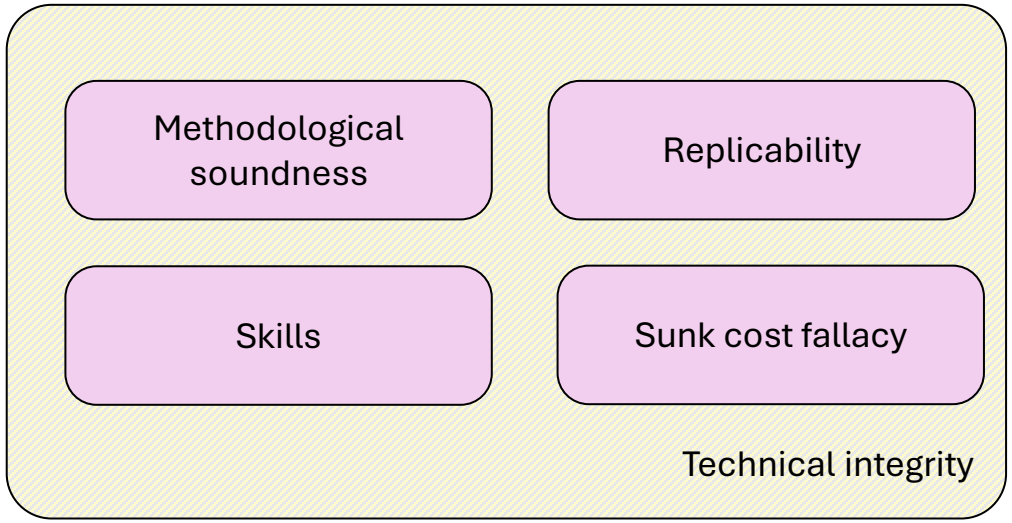
Transparency and Replicability Concerns

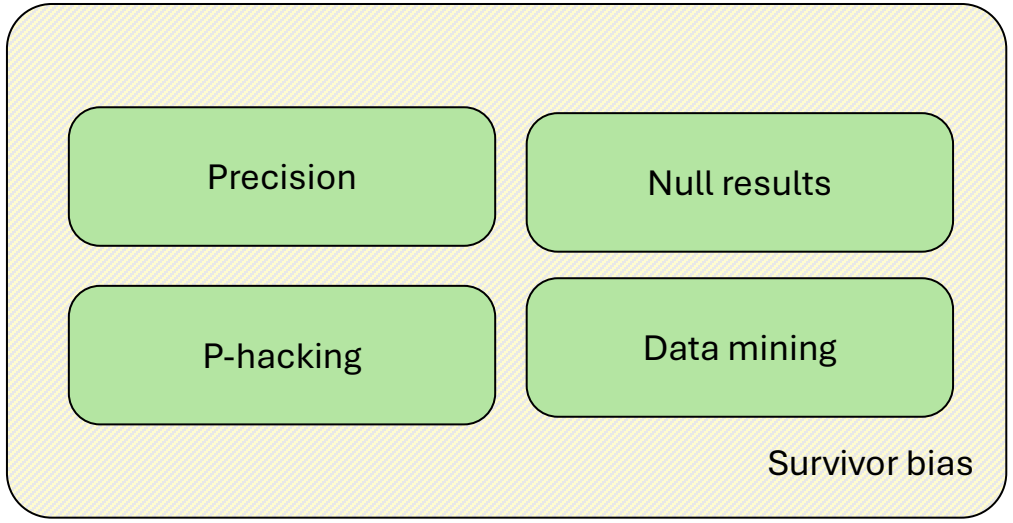
•**Decline in Reporting Null Results:** Reporting of null results declined from 15% in 1980 to around 8.6% in 2023, possibly reflecting increased pressure to produce significant findings and contributing to publication bias.

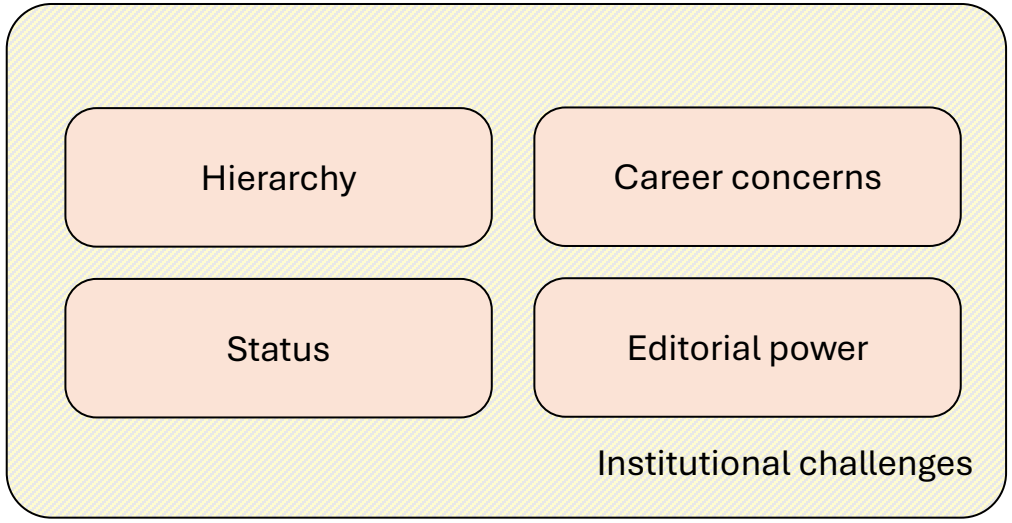
•**Increase in Use of Private Data:** The use of private data doubled from about 4% in 1980 to above 8% in 2023, raising concerns about data accessibility, replicability, and transparency in economic research.

Some of the dimensions we relate to

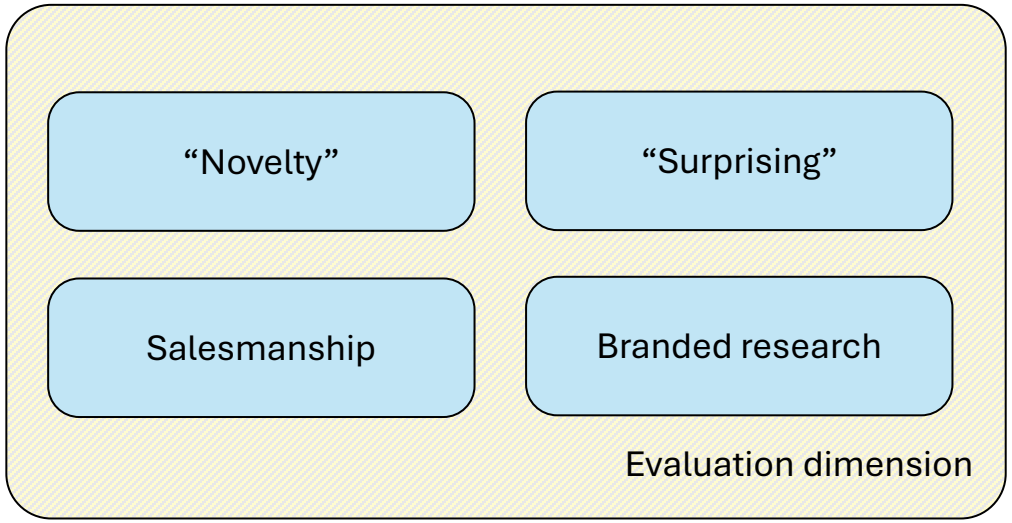








Evaluation of research



- Are "new" ideas getting harder to find?
  (Bloom et al., 2020)
- Are papers becoming less disruptive over time?
  (Park et al., 2023)

- Much of economics research revolves around answering the same old questions with slightly different methods
- Runs the risk that "framing" of research and active salesmanship of research work becomes a much more important vector of dissemination as genuine novelty (Andre et al, 2021).
- Access to "clubs" or networks may be a more important factor driving publication success than actual research quality.

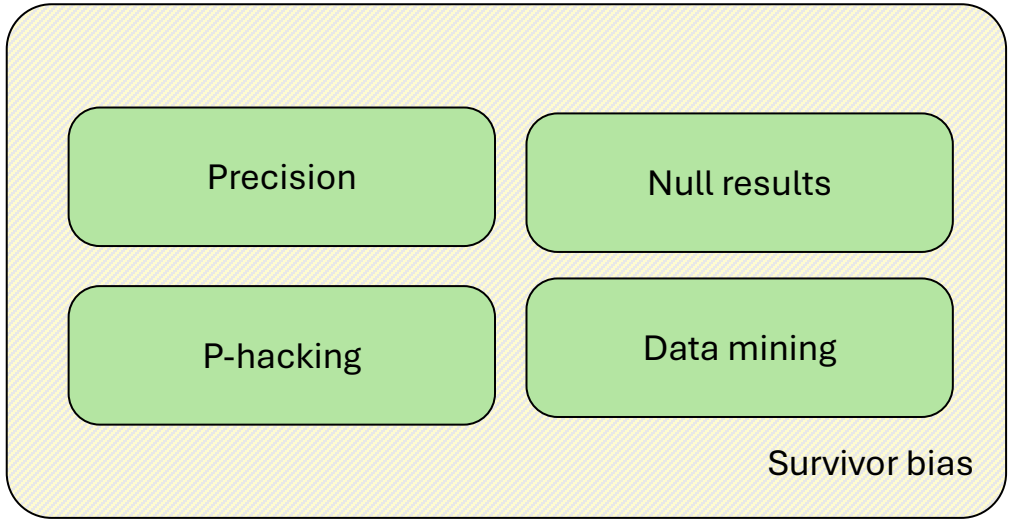
Here: **publication outcomes in Top 5** and attempt to develop a measure of **narrative complexity** versus **overclaiming** or, more colloquially "overselling".

Specifically: what kind of causal narrative do top journals favour?

What constitutes a finding

- No market for “null results” in most journals (Brodeur et al., 2016)

- Limited conceptual or deeper understanding of small effect sizes e.g. due to quality of measurement or aggregation



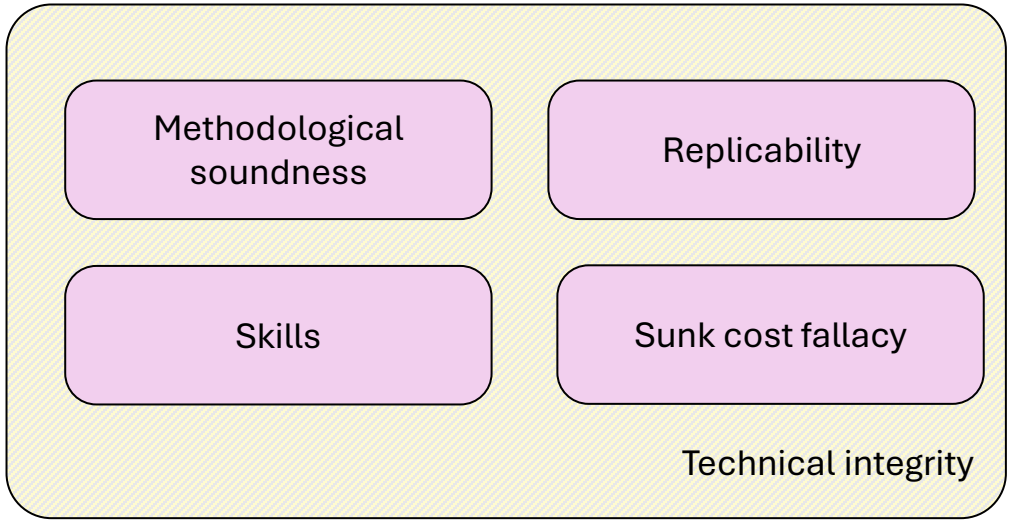
- Data mining is frowned upon, despite mounting evidence of widespread p-hacking suggesting its widespread practice (Simmons et al., 2011; Simonsohn et al., 2014; Gelman and Loken, 2014)

- Loss of comment culture for which journals originally were designed for.

RQ: to what extent economics profession is **comfortable with nulls?**

Research as a skilled trade

- Lack of **quality training** even among established researchers due to technical progress being incredibly fast.
- Best practices on replication are often still only an aspiration not a reality.

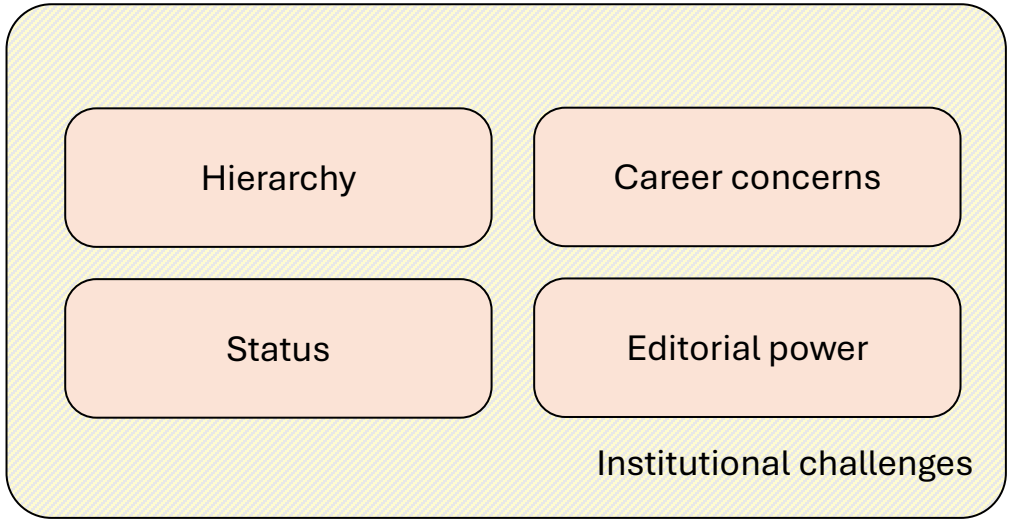


- Focus on cohesiveness within methodological approaches but limited focus on external validity or external relevance (**causal mechanisms** should have **predictive power**)
- **Sunk cost fallacy** may not be irrelevant in some fields of economics.

Here, we wont have much to say at this point.

Incentives shape the process

- Hierarchical organization in economics is an outlier especially in contrast to other fields.
- **Career concern incentives** and limited replication practice have undermined trust.
- Whole fields can be **led down rabbit holes**
- Editorial decisions can make or break a career due to obsession with a few journals.



Here, focus on the publication outcomes along a range of dimensions that may shape incentives.