Risk-taking and innovation in the choice of research topics

A comparative quantitative overview

Carlo Debernardi, Marco Seeber

May 26, INAS 2022

Introduction

General goal:

> Investigating both **contextual** and **individual**-level factors affecting **novelty** contribution in terms of introduction of new research topics

Novelty:

- underinvestigated
- no standard methodology
- recent tools
- usually included as a predictor

Novelty

Novelty as outcome of a recombination process of both old and new pieces of knowledge (March 1991, Schumpeter 1939)

Operationalizations based on *referenced journals*:

- Unprecedented combinations (Wang et al. 2017)
- Atypical combinations (Uzzi et al. 2013)

However these are *input* measures and rather capture *interdisciplinarity* (Fontana et al. 2020)

Growing interest towards text-based measures:
- Finer level of detail
- Focus on semantic content (role of language, e.g. Kuhn 1962)
- Based on the *outcome* of research

Factors affecting novelty

Insights from innovation studies and works focusing on impact:

- Role of newcomers in teams (Katz 1982, Perretti and Negro 2007)
- Redundancy of experience (March 1991)
- Cultural diversity (Jones et al. 2020, Brixy et al. 2020)
- Gender diversity (Campbell et al. 2013)
- Network position, e.g. centrality (Granovetter 1973, Ahuja 2000, Purwitasari et al. 2020)

However, evidence about novelty in science is still limited. Especially if considered as dependent variable rather then a predictor of impact.

Contextual factors

In order to innovate researchers need a *protected space* – i.e. resources, time, stability etc (Whitley 2000)

- **Disciplinary features**
  - *mutual dependence*: level of technicality, reliance on other's work, team/lab work
  - *task uncertainty*: complexity of phenomena, fuzziness of results

- **Country-level contextual factors**
  - different *administrative traditions* (Napoleonic, Germanic, Scandinavian etc; Peters 2021)
  - research evaluation, funding and recruitment policies
  - different levels of NPM policies implementation

Design

Comparative study across:

Countries

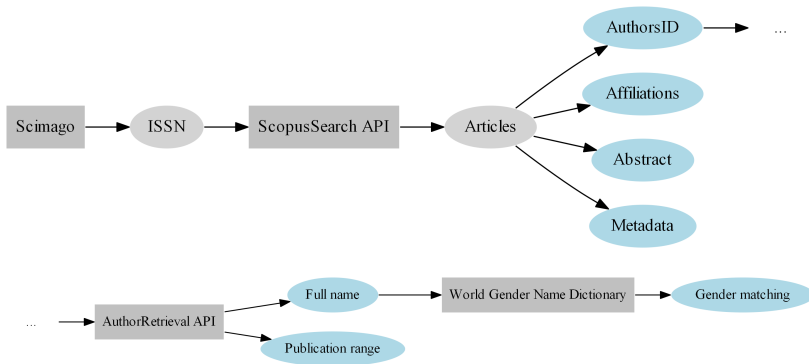- Italy
- Germany
- Netherlands
- Norway

Fields

- Sociology (and Political Science)
- Statistical and Nonlinear Physics
- Molecular Biology

Unit of analysis: paper

Dependent variable: novelty

Predictors: variables aggregated at the article level

Data pipeline for each field (timespan 1990-2020):

Data

	Sociology	Stat and Nonlinear	Molecular Biology
Articles	16,560	28,436	-
m authors	8,172	19,718	-
f authors	6,489	3,381	-
u authors	1,900	5,338	-

Predictors - article level

- #authors
- Publication year
- #juniors
- Seniority
- Gender diversity

- International diversity
- Institutional diversity
- Funded
- Avg degree centrality
- Avg betweenness centrality

Methodological challenges

Measuring novelty
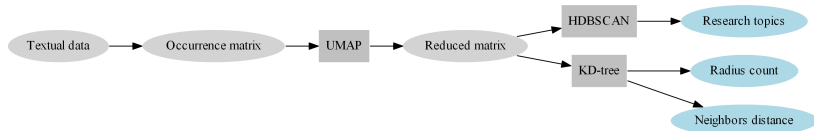
Topic approach

- drastic drop over time (comparability issue)
- simple cut-off hides complexity

Spatial approach

- still dropping over time (standardization, sliding window)
- interpretability issue

Methodological challenges

Topic approach
- Dep variable: dummy, within t years from topic birth
- Model: Logit regression

Spatial approach
- Dep variable: continuous measure
- Model: OLS regression

Low level of explained variance:

- very sparse data
- non-linear relations
- due to article-level aggregation? $\rightarrow$ multilevel modeling

Thanks for the attention

carlodebernardi.github.io
carlo.debernardi.w@gmail.com