Environmental Migration?

A quantitative overview of the literature

Maria Cipollina, Luca De Benedictis, Elisa Scibè

October 10, 2025

5th European Conference on Social Networks, EUSN2021
YoungARS Session
September 8th, 2021

Overview

Background and Motivation

- Increasing scientific evidence that world's climate is changing and related natural hazards are increasing (IPCC Sixth Assessment Report)
- Adverse environmental conditions affect many aspects of human activities, including mobility
- The impact of climate change and natural disasters on migration has gained both public and academic interest in the last decades.
- This relation does not seem straightforward and lead to heterogeneous approaches and empirical findings on the direction of the impact
- Our aim is to map the economic literature analysing the relationship between environmental factors and human migration.
- Develop a methodology that combines, explains, provides a taxonomy and summarises a large number of empirical findings on the topic.

Background and Motivation

The majority of empirical outcomes find an active role of environmental factors ad migration driver.
$\rightarrow$ However, contrasting results are not rare to find.

Reasons may be found in

- Type of environmental event (slow- or fast-onset event) and measurement
- Type of migration (internal or international) and measurement
- Context of the analysis (countries, areas...)
- Characteristics of the area or channels of transmission of the effect (agricultural dependence, conflicts, development, public aid and assistance, etc.)
- Level of analysis (micro or macro)
- Source of data
- Theoretical models
- Empirical strategies

Methodology

Our methodology is developed in three stages:

1. Systematic bibliographic review
2. Citation-based network and community detection
3. Meta-analysis

Systematic Review

Definition of the boundaries of the literature

Environmental factors (slow- and fast-onset events) as driver of migration (international and internal, including urbanisation)

Slow-onset events:

- Temperature
- Precipitation
- Soil degradation

Fast-onset events (disasters):

- Geophysical
- Meteorological
- Climatological
- Hydrological

Systematic Review

Data collection

- Literature database (only published papers)
  - Scopus: 81 records
  - Web of Science: 78 records
- Previous Meta-Analyses
  - Hoffman et al. (2020): 30 records
  - Beine and Jeusette (2019): 51 records
- RePEc (not published papers): 23 records

After duplicates removal: 203 documents

Screening of the results

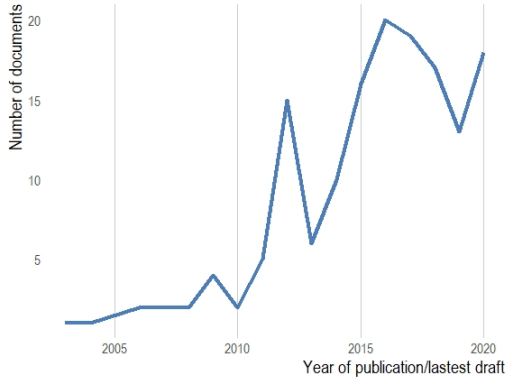
- Screening by title
- Screening by abstract and text

Final sample: 151 documents

Systematic Review - Bibliometric Analysis

Time-span

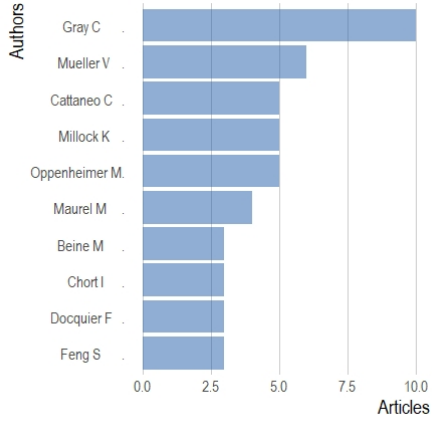
Figure: Time-span of scientific production on environmental migration



Systematic Review - Bibliometric analysis

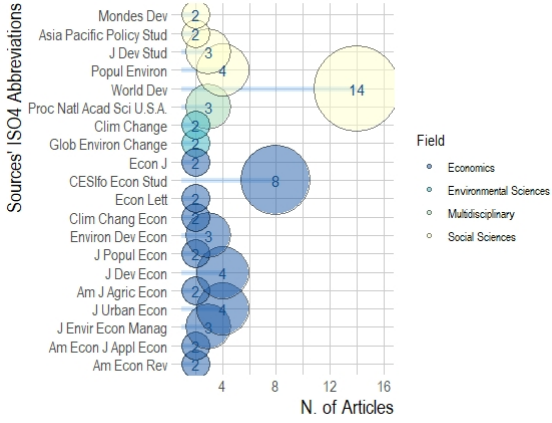
Authors

Description	Results
Authors	288
Author Appearances	372
Female	*150 (40.3%)*
Male	*222 (59.7%)*
First author	
Female	*62 (41%)*
Male	*89 (59%)*
Single-authored documents	34
Authors per Document	1.88
Collaboration Index	2.16



Systematic Review - Bibliometric Analysis

Sources



Bibliographic Coupling Network

To study the inter-connectivity of papers, we build a citation-based network of all 151 papers of our sample

Bibliographic Coupling

Two scientific contributions *bear a meaningful relation to each other when they have one or more references in common* (Kessler, 1963).
A reference that is cited by two papers constitutes a *unit of coupling between them*
Two articles are then said bibliographically coupled if at least one cited document appears in both bibliographic references

Bibliographic Coupling Network

We start from a bipartite network, a rectangular binary matrix that links each paper to its references.

$$\mathrm{A} = \textit{Papers} \times \textit{References} \tag{1}$$

- A is a matrix 151×5.433
- element a_{ij} equals 1 when paper i lists paper j in its bibliography; 0 otherwise

To obtain the *bibliographic coupling* network

$$\mathrm{B} = \mathrm{A} \times \mathrm{A}^T \tag{2}$$

where A is the bipartite network and A^T is its transpose (Aria & Cuccurullo, 2017)

Bibliographic Coupling Network

Matrix B is a symmetrical matrix 151 $\times$ 151 where:

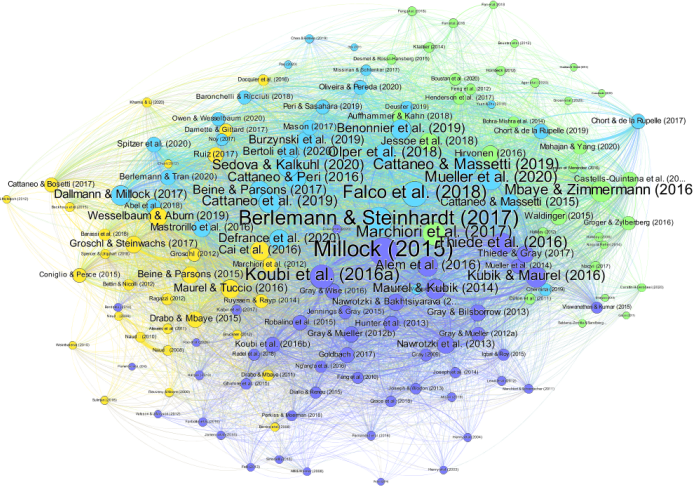
- rows and columns represent a paper included in the sample
- two papers are bibliographically coupled if the value of the tie between them is not zero, meaning they do not share any common reference
- two bibliographically coupled papers are linked by the number of shared common references in their bibliographies.

Therefore, we obtain a *undirected weighted network*

- nodes: 151
- edges: 6063

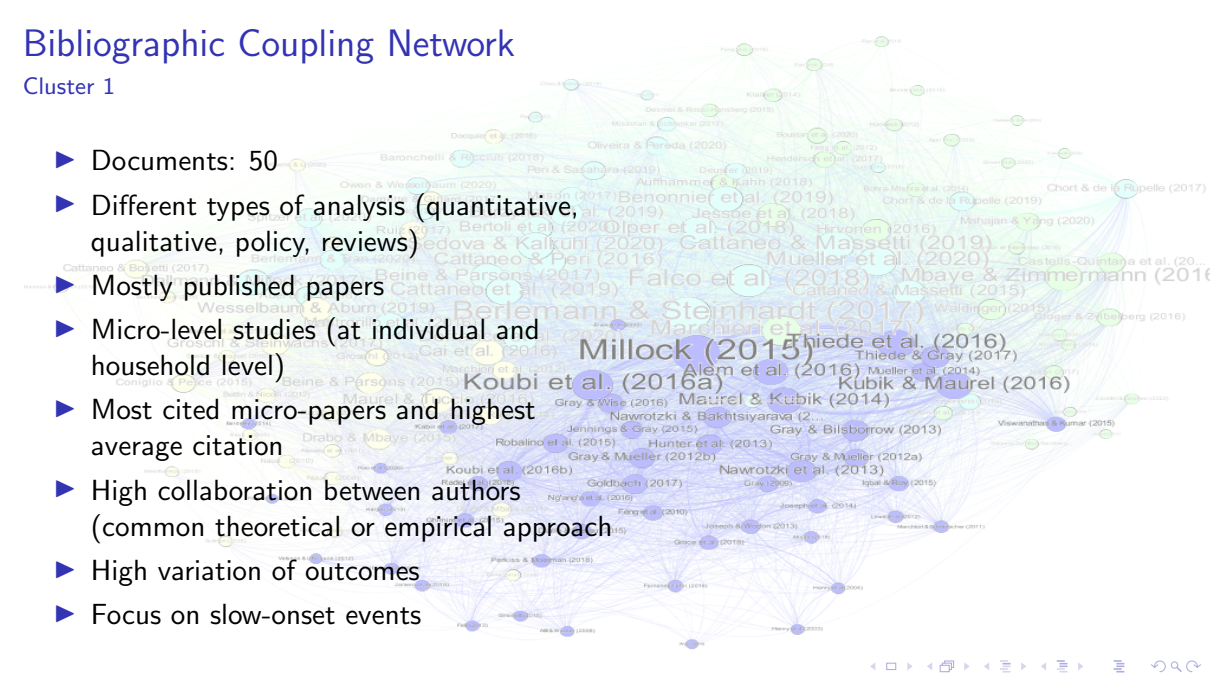
Bibliographic Coupling Network

Community Detection



Bibliographic Coupling Network

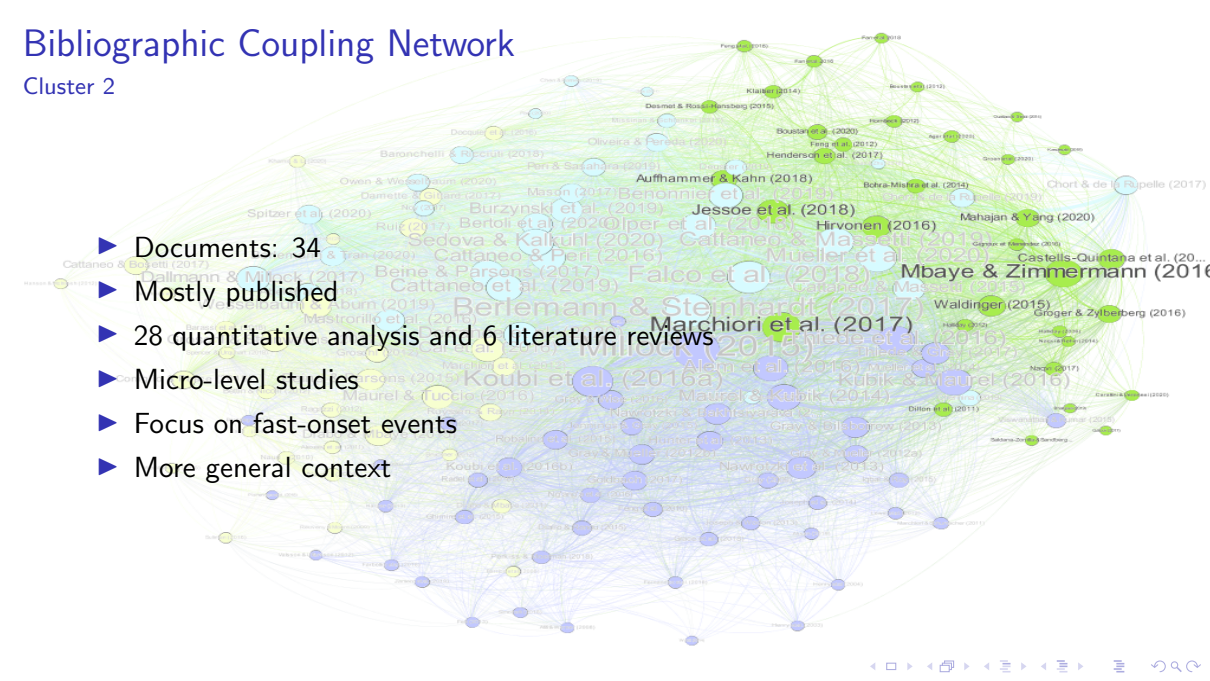
Cluster 1



- Documents: 50
- Different types of analysis (quantitative, qualitative, policy, reviews)
- Mostly published papers
- Micro-level studies (at individual and household level)
- Most cited micro-papers and highest average citation
- High collaboration between authors (common theoretical or empirical approach
- High variation of outcomes
- Focus on slow-onset events

Bibliographic Coupling Network

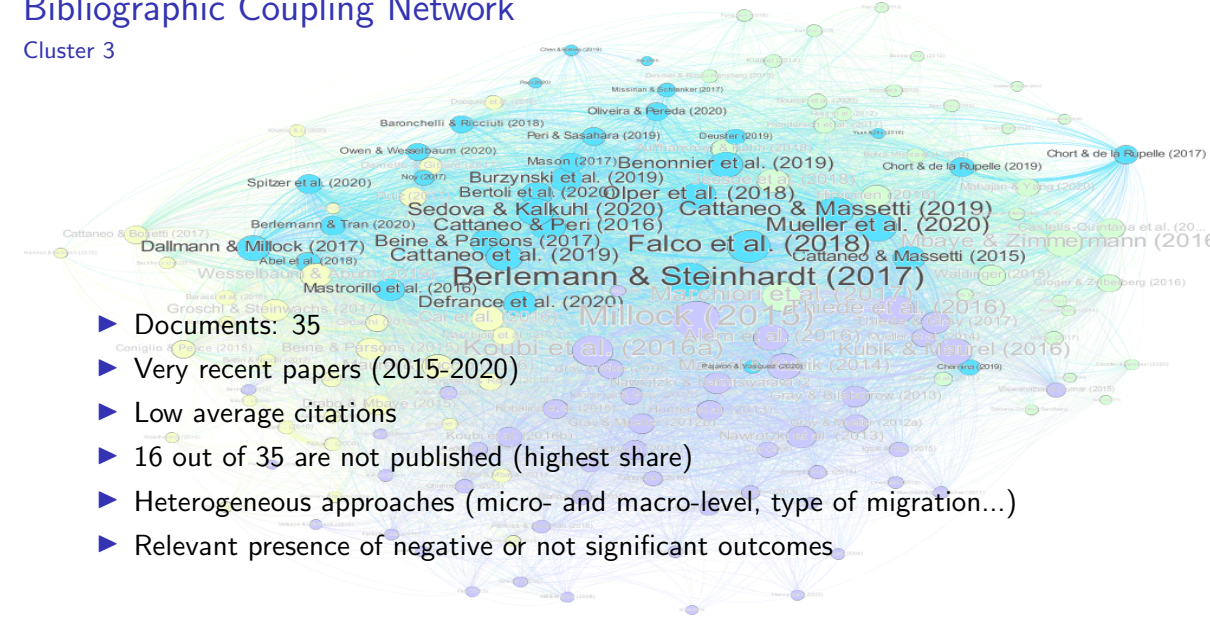
Cluster 2



- Documents: 34
- Mostly published
- 28 quantitative analysis and 6 literature reviews
- Micro-level studies
- Focus on fast-onset events
- More general context

Bibliographic Coupling Network

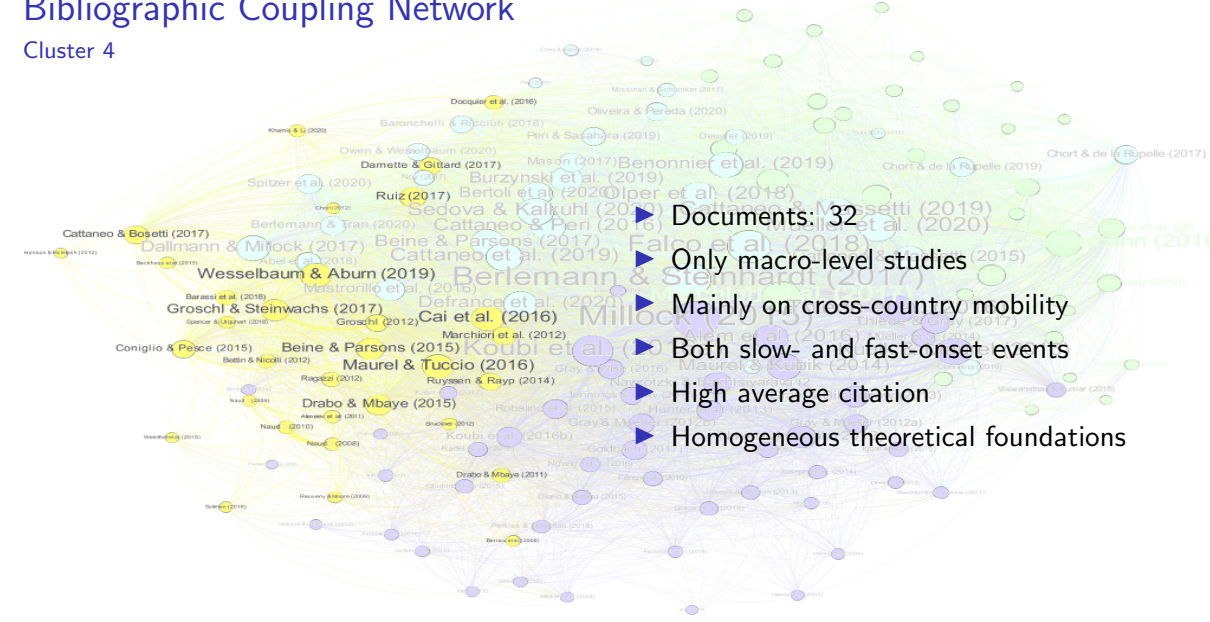
Cluster 3



- Documents: 35
- Very recent papers (2015-2020)
- Low average citations
- 16 out of 35 are not published (highest share)
- Heterogeneous approaches (micro- and macro-level, type of migration...)
- Relevant presence of negative or not significant outcomes

Bibliographic Coupling Network

Cluster 4



- Documents: 32
- Only macro-level studies
- Mainly on cross-country mobility
- Both slow- and fast-onset events
- High average citation
- Homogeneous theoretical foundations

Meta-Analysis

Meta-analysis is a "quantitative survey" of a literature that reports econometric estimates of the same parameter of interest.

It requires that the studies treated are so similar that their differences can be codified.

It allows to assemble the results of multiple trials of similar treatment into a single cumulative average effect

Meta-Analysis

Coded variables

For each regression included in a paper we coded variables to summarise that account for potential sources of heterogeneity:

- **Regression characteristics** (preferred specification/robustness check, estimation strategy)
- **Data characteristics** (source, level of observation, time span, structure of data)
- Value and characteristics of each estimated **coefficients** plus additional info (standard errors, interaction terms, n. of observation, measure of fit, details about standard errors)
- **Dependent variable** (type of mobility - cross-country or internal, included urbanisation -, measure of mobility - flows, rate, stock, dummy -, origin and destination areas)
- **Independent variable**
  - Slow-onset events (measure of temperature and precipitation - level, variation, anomaly -, time lag, soil degradation)
  - Fast-onset events (measure of natural disasters - occurrence, frequency, intensity, duration, losses -, time lag, type of disaster - geophysical, meteorological, hydrological, climatological and subgroups)
- Inclusion of **control variables** and channels (income, conflict, agriculture, etc.)

Meta-Analysis

Dataset

Final sample of 96 papers, all containing at least one or more equation which estimate the impact of environmental factors on different left-hand side variables used as proxy of migration

- 3.904 observations of point estimates of slow-onset events
- 2.065 observations of point estimates of fast-onset events

Meta-Analysis

Standardisation of coefficients

To compare estimates and correctly interpret the synthetic results, a standardisation is needed.

Here, estimates from separate, but similar studies, are converted in **partial correlation coefficients** (*pcc*), commonly used in MA literature (Doucouliagos, 2005; Stanley and Doucouliagos, 2012).

Partial Correlation Coefficients

The *pcc* allow to analyse within a single framework all available studies on the effects of climate changes on migration regardless of the specification or measure of migration used. $r_i = \frac{t_i}{\sqrt{t_i^2 + df_i}}$

$$se_i = \sqrt{\frac{(1 - r_i^2)}{df_i}}$$

Meta-Analysis

Standardisation of coefficients

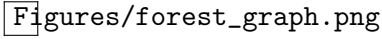


Figure: Box Plot of Partial Correlation Coefficients

Meta-Analysis

Basic MA - slow-onset events

	Model	Averages		I^2	Q-test (p-value)
Slow onset effect	FEM	0.0001	***	86.78	0.00
	REM	0.0006		99.93	0.00
Cluster 1	FEM	0.0001	***	83.15	0.00
	REM	-0.0025	**	99.97	0.00
Cluster 2	FEM	0.0003		95.32	0.00
	REM	0.0068		99.84	0.00
Cluster 3	FEM	-0.0037	***	77.58	0.00
	REM	-0.0039	***	93.58	0.00
Cluster 4	FEM	0.0060	***	88.44	0.00
	REM	0.0082	***	94.96	0.00

Meta-Analysis

Basic MA - fast-onset events

	Model	Averages		I^2	Q-test (p-value)
Fast onset effect	FEM	0.0021	***	91.42	0.00
	REM	0.0085	***	97.76	0.00
Cluster 1	FEM	0.0022	***	86.50	0.00
	REM	0.0140	***	98.98	0.00
Cluster 2	FEM	-0.0021	***	85.84	0.00
	REM	-0.0033		98.77	0.00
Cluster 3	FEM	-0.0004		80.19	0.00
	REM	0.0028	***	89.04	0.00
Cluster 4	FEM	0.0071	***	96.11	0.00
	REM	0.0224	***	98.94	0.00

Meta-Regression

$$r_i = \beta_0 + \beta_1 se_i + \epsilon_i$$

Weighted Least Squares

WLS corrects the previous equation for heteroscedasticity

$$t_i = \frac{r_i}{se_i} = \beta_1 + \beta_0 \frac{1}{se_i} + \epsilon_i$$

We then run **5 meta-regressions** for each type of environmental stressor (full sample and cluster by cluster) adding all coded controls

Meta-Analysis

Results and Conclusions

- Overall, the meta-analytic average effect estimated through the meta-regression shows a positive and significant effect of slow-onset events on migration. However very small in magnitude
- The average effect of fast-onset events is positive but not significant.
- When all controls are added it is indeed clear that all sources of heterogeneity in each study create biases in the estimated outcomes
- When the communities are accounted for and we run separate meta-regressions, a significant heterogeneity emerges. Signs, magnitudes and statistical significance are consistently different
- Evidence on the formation of club-like convergence of literature outcomes

Thank you!