Modelling growth of urban firm networks

J. Raimbault[1,2,3,4,*] N. Zdanowska[5,2,4] E. Arcaute[2]

*juste.raimbault@ign.fr

[1]LASTIG, Univ Gustave Eiffel, IGN-ENSG
[2]CASA, UCL
[3]UPS CNRS 3611 ISC-PIF
[4]UMR CNRS 8504 Géographie-cités
[5]Luxembourg Institute of Socio-Economic Research

CCS 2022
UrbanSys Satellite
19/10/2022

Urban firm networks

Role of firm networks in urban systems:

- Cities as cross-overs of socio-economic interactions [Castells, 1996]
- Networks as the new social morphology of societies [Castells, 2000], giving rise to network economies [Sassen, 1991]
- Emergence of an interconnected world city network as a crucial feature of globalisation [Taylor, 2001]
- Interurban interactions between firms as one of the most representative of economic trends [Martinus and Sigler, 2018]

An evolutionary approach to Urban Systems

Relevance of dynamics urban network evolution models:

- Evolutionary Theory of Urban Systems [Pumain, 1997] [Pumain, 2006]
- Adaptive cycles and diffusion of innovation [Hagerstrand, 1968]
- Path dependence [Martin and Sunley, 2006] [Pumain, 2012]
- Selection and emerging structures of systems
- Evolutionary models for urban systems dynamics: [Favaro and Pumain, 2011] [Cottineau et al., 2015] [Schmitt et al., 2015] [Raimbault, 2018] [Raimbault, 2020] [Raimbault et al., 2020]

Urban firm linkages and geo-economic processes

Main geographical processes considered:

- Metropolisation & internationalisation effects
- Specialisation-driven factors and drivers of innovation
- Macroeconomic exogenous chocs and resilience of urban systems

$\rightarrow$ *How can we capture geographical and economic processes within urban networks of firms with a generative model?*

Advantages of the model introduced:

- Indirect inference on processes
- Multiple complementary factors driving inter-urban links
- Includes path-dependency

Contribution

We examine the interactions of European cities within firm linkages defined by corporate ownership links and introduce

- an empirical analysis of the firm ownership urban network in the European Union, based on the Bureau Van Dijk's AMADEUS database;
- a generative network model to simulate the growth of such linkages at urban areas scale;
- the model allows us to compare the effect of different factors on the final network structure, which is extensively studied through model sensitivity analysis and exploration;
- we calibrate the model on the empirical network and apply it to stylized economic shock scenarios.

Model rationale

$\rightarrow$ Cities are defined by their GDP and their profile regarding the proportion of firms in the different sectors. Links between cities are created in an iterative way, taking into account:

- geographical proximity (distance or effective accessibility)
- geopolitical proximity (belonging to the same country or single market $\rightarrow$ application to Brexit)
- city size (economic size as GDP)
- economic similarity (e.g. cosine distance between sector proximity as done in [Cottineau and Arcaute, 2019]
- previous linkages

Formalization

Cities characterised by economic size E_i (GDP) and economic structure S_{ik} (probability distribution of firms within K sectors)

Starting from an initial network, at each time step, add a fixed number of links randomly, following a probability as a generalised Cobb-Douglas function

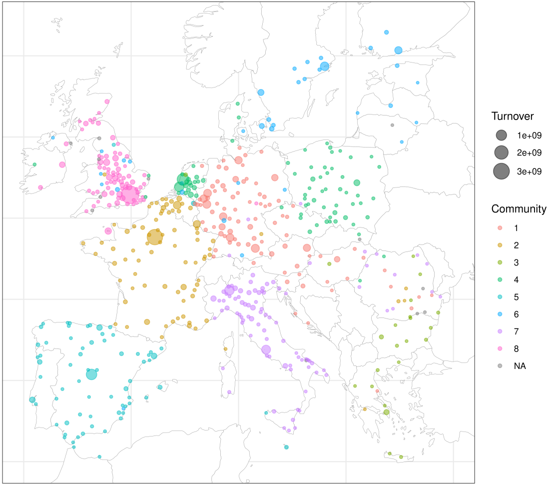
$$p_{ij} \propto \left(\frac{E_i}{E}\right)^{\gamma_O} \cdot \left(\frac{E_j}{E}\right)^{\gamma_D} \cdot \left(\frac{w_{ij}}{W}\right)^{\gamma_W} \cdot s\left(S_{ik}, S_{jk}\right)^{\gamma_S} \cdot \exp\left(-\frac{d_{ij}}{d_0}\right) \cdot \exp\left(-\frac{c_{ij}}{c_0}\right)$$

where $E = \sum_k E_k$, $W = \sum_{i,j} w_{ij}$, s is a proximity measure given by cosine similarity, d_{ij} euclidian distance, and c_{ij} a socio-cultural distance

Model parameters: $\gamma_O, \gamma_D, \gamma_W, \gamma_S, d_0, c_0$

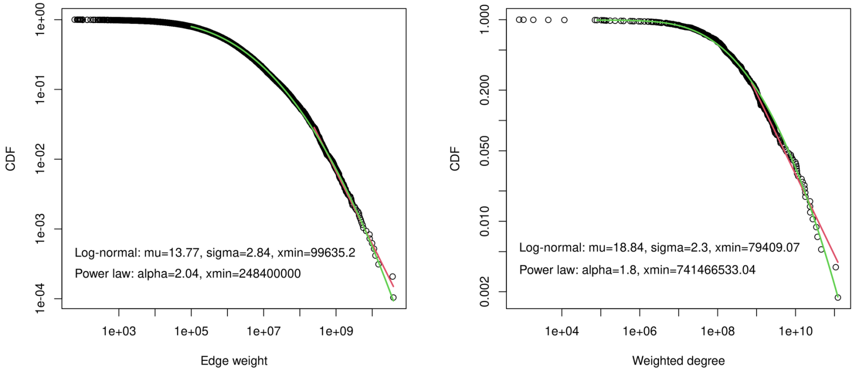
Model indicators: internationalisation (modularity of countries in the network), metropolisation (correlation between weighted degree and city size), optimal modularity, average community size

Empirical data



Map of network communities obtained by modularity maximisation (European ownership network between Functional Urban Areas, constructed from the AMADEUS database).

Empirical network properties



Empirical distribution of network properties, with log-normal and power law fits

Statistical models

Most general spatial interaction model with fixed effects:

$$\log w_{ij} = \log d_{ij} + \log T_i + \log T_j + \log s_{ij} + \alpha_{c_i,c_j} + \varepsilon_{ij} \tag{1}$$

	d_0		c_0		γ_S		γ_W		γ_O		γ_D	
	F	T	F	T	F	T	F	T	F	T	F	T
Internationalisation	0.2	0.3	0.7	0.7	0.001	0.008	$5 \cdot 10^{-4}$	0.007	0.03	0.04	0.02	0.04
Metropolisation	0.02	0.04	0.02	0.04	0.02	0.3	0.002	0.01	0.2	0.7	0.2	0.7
Modularity	0.3	0.4	0.6	0.6	$5 \cdot 10^{-4}$	0.02	0.002	0.02	0.007	0.03	0.003	0.03
Avg. com. size	0.004	0.1	0.01	0.1	0.003	0.07	0.001	0.04	0.3	0.6	0.3	0.6
Degree entropy	0.003	0.03	0.002	0.03	0.008	0.04	0.003	0.03	0.5	0.6	0.5	0.6
Weight entropy	0.04	0.2	0.03	0.2	0.002	0.1	0.001	0.1	0.5	0.6	0.5	0.6

Zero-inflated and Hurdle models: Details

Model setup

1. Simulation on synthetic systems of cities [Raimbault et al., 2019]: generation of a continent-scale urban system with stylized order of magnitude corresponding to Europe (urban hierarchy, countries, industrial sector distributions - Details)
2. Simulation on the European system of cities, setup with AMADEUS data at FUA level (Details)

Implementation and experiments

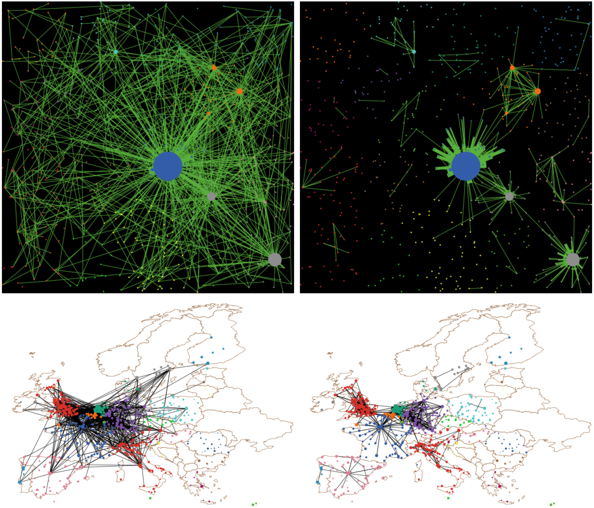
Implementation

- Model implemented in NetLogo (good compromise interactivity / ergonomy), with fast data structures (matrix/table extensions)
- Integrated seamlessly into OpenMOLE [Reuillon et al., 2013] for model exploration `https://openmole.org/`

Experiments

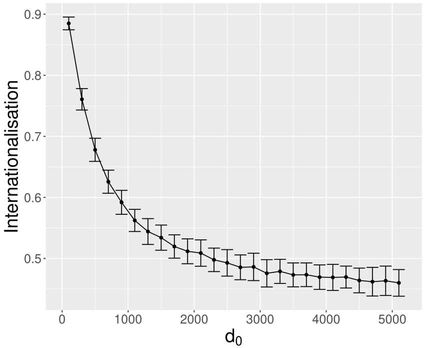
- One-factor sampling with 100 repetitions to assess statistical properties (good convergence, average sharpe ratios for indicators all larger than 5)
- Grid sampling with 20 repetitions for model behavior
- Targeted experiment on impact of synthetic hierarchy
- Global Sensitivity Analysis [Saltelli et al., 2008]
- Calibration on real setup with a Genetic Algorithm
- Application to economic shocks scenarios

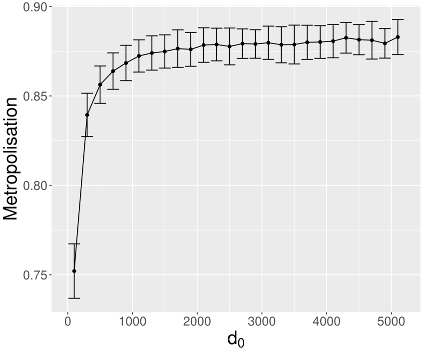
Simulation of urban networks



Simulated networks in a synthetic setup (top) and real setup (bottom), for a large (resp. low) interaction range d_0 (left, resp. right).

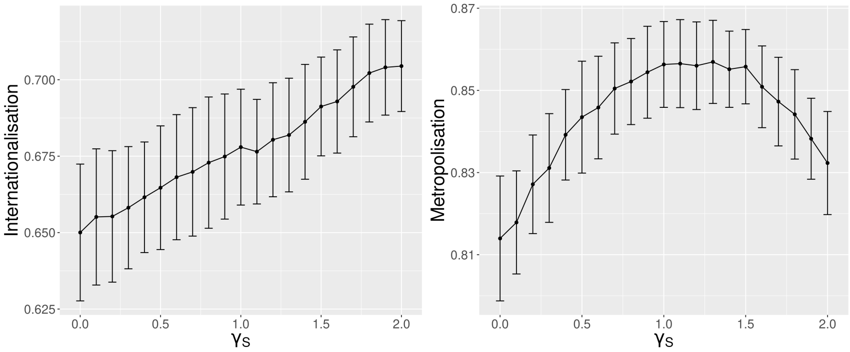
Effect of interaction decay





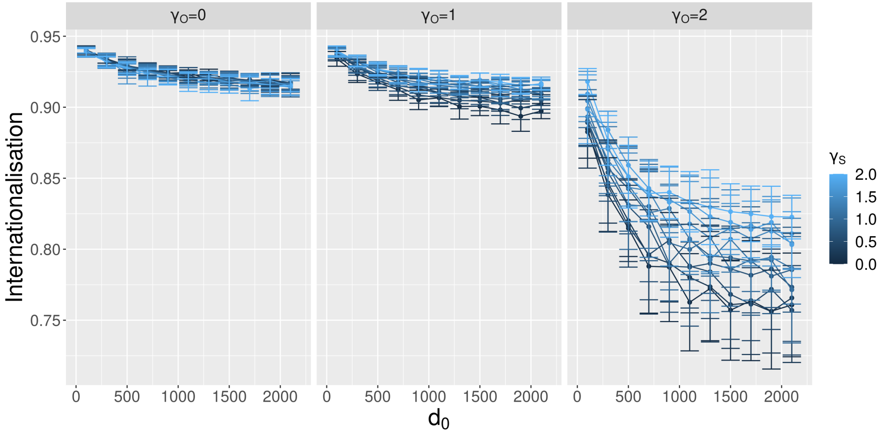
(Left) Internationalisation index decreases exponentially with gravity decay; (Right) metropolisation also shows a progressive transition from a local to a global regime.

Effect of sector proximity



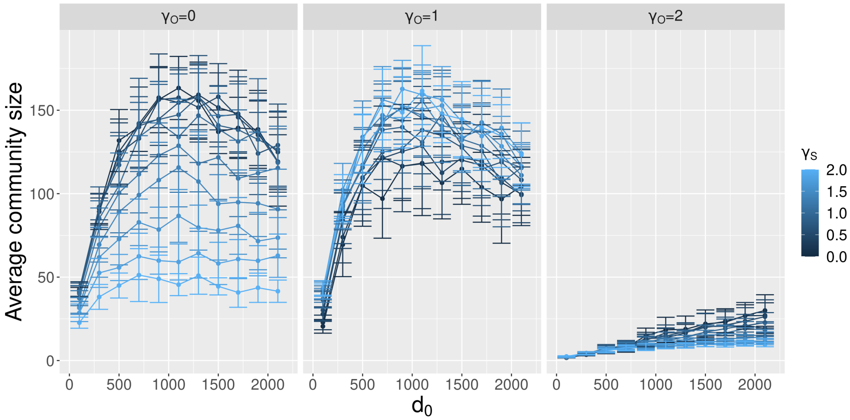
(Left) Internationalisation varies linearly with sector proximity γ_S; (Right) metropolisation witnesses an intermediate regime where size is the most important

Grid exploration: internationalisation



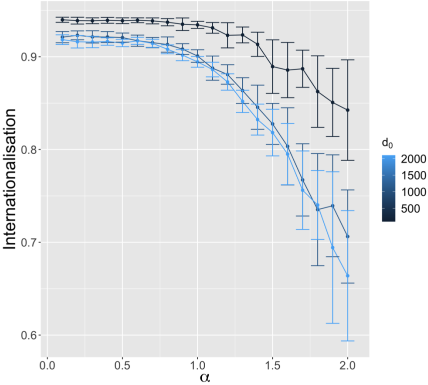
The transition as a function of interaction range depends on the influence of origin size γ_F; sector proximity γ_S plays a role only for a large influence of the origin.

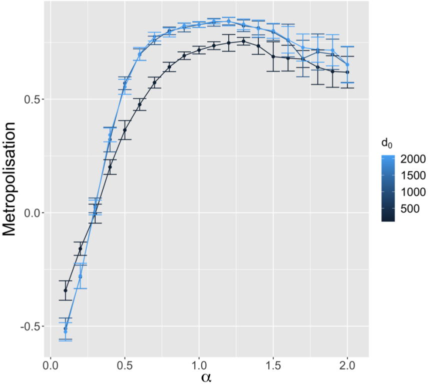
Grid exploration: community size



- Maximal integration in term of community size is achieved at an intermediate value of d_G: emergence of a regional regime
- Maximal size depends on the role of sectors γ_S, in a decreasing way when origin size is deactivated, and increasing way when $\gamma_F = 1$
- This regime disappear when origin size influence is too large

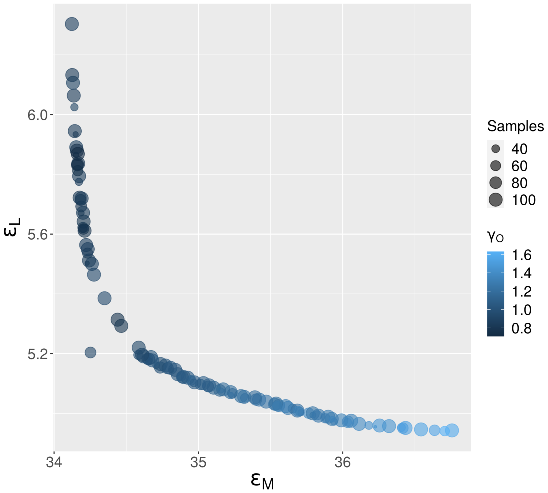
Role of urban hierarchy





Impact of urban hierarchy on model indicators in synthetic systems of cities: existence of a maximum for metropolisation at intermediate hierarchies.

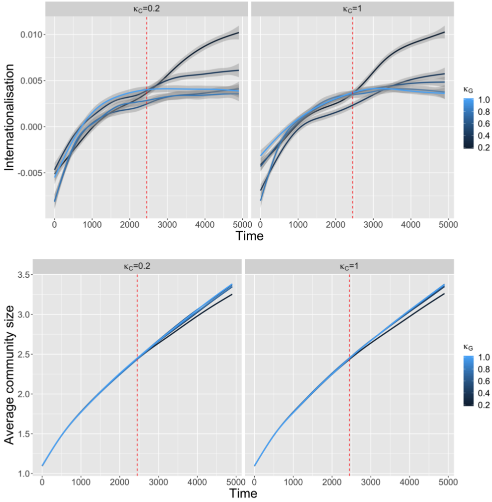
Calibration on the European urban network



Pareto front obtained for the bi-objective model calibration using a NSGA2 algorithm (outperforms statistical models that have a MSE of 5.33).

Details on the calibration procedure

Economic shock scenarios



Although the system is resilient to moderate shocks, strong restrictions have an important impact on internationalisation and thus can be expected to have a detrimental effect on UK economy due to these foreign ownership links.

Discussion

Practical applications

→ Potential application to public policy issues;

→ Effect of exogenous shocks on specific economic sectors mostly dependent on foreign ownership links (extraction of crude petroleum, manufacture of tobacco & basic pharmaceutical goods, activities of head offices).

Developments

→ co-evolution model (evolution of city sizes) [Raimbault, 2021];

→ a model with firm agents (multi-scale ABM) [Raimbault, 2019];

→ targeted study of path dependency;

→ other formulation of the combination of factors or multi-objective optimisation depending on sectors using Pareto fronts.

Conclusion

→ A generative model to understand processes of economic network emergence

→ Crucial role of model exploration to validate and extract knowledge from such a simulation model

Preprint at `https://arxiv.org/abs/2009.05528`

Open repository for model and results at `https://github.com/JusteRaimbault/ABMCitiesFirms`

Simulation data at `https://doi.org/10.7910/DVN/UPX23S`

Acknowledgments: thanks to the *European Grid Infrastructure* for access to the infrastructure. We also acknowledge the funding of the EPSRC grant EP/M023583/1

References I

Banos, A., Moreno, D., Pivano, C., and Taillandier, P. (2011).
Christaller, toujours vivant!
Cybergeo: European Journal of Geography.

Castells, M. (1996).
The information age: Economy, society and culture (3 volumes).
Blackwell, Oxford, 1997:1998.

Castells, M. (2000).
The rise of the network society.
Blackwell Publishers, Cambridge MA.

Cottineau, C. and Arcaute, E. (2019).
The nested structure of urban business clusters.
arXiv e-prints, page arXiv:1905.05106.

References II

Cottineau, C., Chapron, P., and Reuillon, R. (2015).
Growing models from the bottom up. an evaluation-based incremental modelling method (ebimm) applied to the simulation of systems of cities.
Journal of Artificial Societies and Social Simulation, 18(4):9.

Favaro, J.-M. and Pumain, D. (2011).
Gibrat revisited: An urban growth model incorporating spatial interaction and innovation cycles.
Geographical Analysis, 43(3):261–286.

Hagerstrand, T. (1968).
Innovation diffusion as a spatial process.
Chicago, University of Chicago Press, page 334.

Martin, R. and Sunley, P. (2006).
Path dependence and regional economic evolution.
Journal of Economic Geography, 6–4:395–473.

References III

Martinus, K. and Sigler, T. J. (2018).
Global city clusters: theorizing spatial and non-spatial proximity in inter-urban firm networks.
Regional studies, 52(8):1041–1052.

Pumain, D. (1997).
Pour une théorie évolutive des villes.
L'Espace géographique, pages 119–134.

Pumain, D. (2006).
Villes et systèmes de villes dans l'économie.
Revue d'économie financière, pages 29–46.

Pumain, D. (2012).
Urban systems dynamics, urban growth and scaling laws: The question of ergodicity.
In *Complexity theories of cities have come of age*, pages 91–103. Springer.

Raimbault, J. (2018).
Calibration of a density-based model of urban morphogenesis.
PloS one, 13(9):e0203516.

Raimbault, J. (2019).
A multi-scalar model for system of cities.
In *Conference on Complex Systems 2019*, Singapore, Singapore.

Raimbault, J. (2020).
Indirect evidence of network effects in a system of cities.
Environment and Planning B: Urban Analytics and City Science, 47(1):138–155.

Raimbault, J. (2021).
Modeling the co-evolution of cities and networks.
In *Handbook of Cities and Networks*, pages 166–193. Edward Elgar Publishing.

References V

Raimbault, J., Cottineau, C., Le Texier, M., Le Nechet, F., and Reuillon, R. (2019).
Space matters: Extending sensitivity analysis to initial spatial conditions in geosimulation models.
Journal of Artificial Societies and Social Simulation, 22(4).

Raimbault, J., Denis, E., and Pumain, D. (2020).
Empowering urban governance through urban science: Multi-scale dynamics of urban systems worldwide.
Sustainability, 12(15):5954.

Reuillon, R., Leclaire, M., and Rey-Coyrehourcq, S. (2013).
Openmole, a workflow engine specifically tailored for the distributed exploration of simulation models.
Future Generation Computer Systems, 29(8):1981–1990.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010).
Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index.
Computer Physics Communications, 181(2):259–270.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008).
Global sensitivity analysis: the primer.
John Wiley & Sons.

Sassen, S. (1991).
The global city.
New York.

References VII

Schmitt, C., Rey-Coyrehourcq, S., Reuillon, R., and Pumain, D. (2015).
Half a billion simulations: Evolutionary algorithms and distributed computing for calibrating the simpoplocal geographical model.
Environment and Planning B: Planning and Design, 42(2):300–315.

Simini, F. and James, C. (2019).
Testing heaps law for cities using administrative and gridded population data sets.
EPJ Data Science, 8(1):1–13.

Taylor, P. J. (2001).
Specification of the world city network.
Geographical analysis, 33(2):181–194.

Reserve slides

Synthetic setup

1. Generate $N = 700$ cities with size following a power law $E_i = E_0 \cdot i^{-\alpha}$ with $E_0 = 10^{11}$ and $\alpha = 1.1$ (computed on Europe for GDP with cities larger than 50.000 inhabitants)
2. Distribute them randomly in space ([Simini and James, 2019] vs [Banos et al., 2011])
3. Create countries with k-means clustering ($C = 30$)
4. Distribute sectors such that (i) smaller cities are more specialised and (ii) larger cities are more knowledge-based, with a one dimensional axis to position sectors $1/K \ldots 1$ where the density $f(k)$ follows a log-normal with (μ, σ) such that $\sqrt{\mathrm{Var}\, f} = K/2$ for the largest, $\sqrt{\mathrm{Var}\, f} = 1/K$ for the smallest

Back

Real configuration setup

Urban areas (FUAs) positions, economic size initialised from the GHSL database, sector composition from the Amadeus database using classification of firms and aggregating

Initial dummy network ($w_0 = 1$)

Socio-cultural distance c_{ij}: fixed effects coefficients

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
-2.327	1.440	2.451	2.538	3.447	11.797	339

Back

Global sensitivity analysis

	d_0		c_0		γ_S		γ_W		γ_O		γ_D	
	F	T	F	T	F	T	F	T	F	T	F	T
Internationalisation	0.2	0.3	0.7	0.7	0.001	0.009	$4 \cdot 10^{-4}$	0.007	0.03	0.04	0.02	0.04
Metropolisation	0.02	0.1	0.02	0.2	0.002	0.1	0.001	0.09	0.2	0.6	0.3	0.6
Modularity	0.3	0.4	0.6	0.6	0.004	0.02	$3 \cdot 10^{-4}$	0.01	0.005	0.03	0.002	0.03
Avg. com. size	0.008	0.09	0.01	0.1	0.002	0.07	0.003	0.04	0.3	0.6	0.4	0.6
Degree entropy	0.006	0.02	0.003	0.02	0.006	0.03	0.008	0.02	0.5	0.5	0.5	0.5
Weight entropy	0.04	0.1	0.03	0.1	0.008	0.08	0.01	0.07	0.4	0.5	0.4	0.5

First order and total sensitivity indices [Saltelli et al., 2010], for all parameters and indicators.

Zero-inflated and Hurdle statistical models

Accounting for zeros with specific models: same qualitative results

Model	Zero-inflated		Hurdle	
	Count	Zero-infl.	Count	Zero-hurdle
$\log(d_{ij})$	-0.28*** (7e-4)	0.88*** (0.01)	-0.29*** (8e-6)	-0.75*** (0.01)
$\log(T_i)$	0.74*** (4e-6)	-0.74*** (7e-3)	0.74*** (5e-6)	0.67*** (5e-3)
$\log(T_j)$	0.64*** (4e-6)	-0.52*** (7e-3)	0.64*** (4e-6)	0.48*** (5e-3)
$\log(s_{ij})$	0.46*** (2e-5)	-0.33*** (3e-2)	0.45*** (2e-5)	0.27*** (0.02)
R^2	0.1607		0.1661	
AIC	5.85e10		5.85e10	

Back

Model calibration

To avoid overfitting: $c_0 = 0$ to deactivate the mechanism

Calibration objectives:

$$\varepsilon_L = \frac{1}{N^2} \sum_{i,j} \left(\log w_{ij} - \log \hat{w}_{ij} \right)^2$$

$$\varepsilon_M = \log \left(\frac{1}{N^2} \sum_{i,j} \left(w_{ij} - \hat{w}_{ij} \right)^2 \right)$$

On macro indicators? Too much solutions, equifinality

Rescale flows: w_0 is arbitrary, so is optimised with rescaling:

$$\varepsilon_L(k_0) = \frac{1}{N^2} \sum_{i,j} \left(\log w_{ij} - \log \left(k_0 \cdot \hat{w}_{ij} \right) \right)^2$$

what yields

$$k_0 = \exp \left[\frac{1}{N^2} \sum_{i,j} \log w_{ij} - \frac{1}{N^2} \sum_{i,j} \log \hat{w}_{ij} \right]$$